Psychological Bulletin

HARRY HELSON, Editor
Kansas State University

CONTENTS

Published Bimonthly by the
American Psychological Association

Vol. 59, No. 3 MAY 1962

Consulting Editors

O. McNemar
Stanford University

L. J. Postman
University of California, Berkeley

J. R. Rotter
Ohio State University

S. B. Sells
Texas Christian University

W. A. Wilson, Jr.
Depauw University

The ... contains evaluative reviews of research literature ... methodology and instrumentation in psychology. This Journal ... of original research or original theoretical articles.

... with the January 1964 issue, all articles will be preceded by ... words typed on a separate sheet of paper. The abstract should ... *Psychological Abstracts*. Detailed instructions for preparation ... obtained from the Editor or from the APA Central Office ... sent to the Editor, Harry Helson, Department of Psychol... University, Manhattan, Kansas.

... *for publication*. Authors are strongly advised to follow ... in the *Publication Manual of the American Psychological ...* Special attention should be given to the section ... references (pp. 50–60), since this is a particular source ... difficulty ... research literature. *All copy must be double spaced* ... which ... (All manuscripts should be submitted in *duplicate*, one ... original typed copy; author's name should appear only on title ... mimeographed copies are not acceptable and will not be con... ... are prepared for publication; duplicate figures may be sho... ... copies. Authors are cautioned to retain a copy of the ... against loss in the mail and to check carefully the typing of the ... final copy.

... reprints are given to contributors of articles and notes.

... — Virginia Richards
... Manager — *Editorial Assistant*

... including subscriptions, orders of back issues, and changes of ... addressed to the American Psychological Association, 1333 ... Seventeenth ... Washington 6, D.C. Address changes must reach the Subscription ... Office ... of the month to take effect the following month. Undelivered ... copies ... address changes will not be replaced; subscribers should notify ... the post ... will guarantee second class forwarding postage. Other claims ... for undelivered ... must be made within four months of publication.

... $10.00 (Foreign $10.50). Single copies, $2.00.

... Published Bimonthly by
... AN PSYCHOLOGICAL ASSOCIATION, INC.
... Wisconsin Avenue
... and 1333 Sixteenth Street N.W., Washington 6, D.C.

... at Washington, D.C. and at additional mailing offices. Printed in U.S.A.

VOL. 59, NO. 3 MAY 1962

Psychological Bulletin

SOME RESULTS OF COMPARATIVE-PHYSIOLOGICAL INVESTIGATIONS OF HIGHER NERVOUS ACTIVITY[1]

L. G. VORONIN

Institute of Higher Nervous Activity and Neurophysiology of the USSR Academy of Sciences, Moscow

The comparative-physiological trend in the study of the functions of the organism arose in the nineteenth century and developed under the influence of evolutionary views on the origin of animals and man. Comparative physiology of the higher nervous activity is based on the reflex of I. M. Sechenov and I. P. Pavlov, as well as on the evolutionary theory of Charles Darwin. The present article describes the most important results of the comparative-physiological study of higher nervous activity carried out by our laboratory; it also gives an interpretation of these results.[2]

I believe that not only the very essence of our method of investigation will be of interest to the reader, but also the theoretical propositions which underlie the fundamental idea of our research into both common and different features of higher nervous activity in various species of vertebrates. These propositions were formulated in the laboratories of Sechenov and Pavlov and may be reduced in brief to the following:

1. The work of the nervous system, including its higher division—the brain—is effected reflexly. It can be made fully known through the development of the reflex theory based on the principle of determinism (according to which there are no causeless reactions in the organism), the structural principle (there are no functions in the organism which are not adjusted to the structure), and the principle of analysis and synthesis (the nervous system is capable of differentiating and integrating all the influences to which it is subjected, as well as the activity of the organism arising in response to these influences).

2. The method of confronting external influences falling on the organism with its corresponding response makes it possible to understand the nervous mechanism involved in this response.

3. The conditioned reflex is a universal mechanism of activity acquired in the course of the organism's individual life. The researcher can investigate the laws of development of this reflex (beginning with its most elementary form) and of its gradual complication in onto- and phylogenesis.

[1] Reprints of this article may be obtained from the Editor of the *Psychological Bulletin*.

[2] We have utilized here part of the experimental material accumulated by my collaborators as well as by myself at the Pavlov Institute of Physiology of the USSR Academy of Sciences (Leningrad) in 1950–54, at the Institute of Higher Nervous Activity and Neurophysiology of the USSR Academy of Sciences in 1957–60, and at the Moscow Lomonosov State University in 1954–60.

4. In the course of evolution of the animal world there took place only a quantitative growth or complication of higher nervous activity. The latter acquired qualitative distinctions as a result of the transition of the organism to the specifically human mode of interaction with the surrounding medium.

All these propositions undoubtedly require further development, especially in connection with the progress of the analytical methods of modern physiological investigations and with the bright prospects which they open before researchers.

Comparative physiological investigations of higher nervous activity conducted by our laboratory are by no means unique in the Soviet Union. Considerable and fruitful work in this field is done by the laboratory of D. A. Biryukov (1959, 1960) at the Leningrad Institute of Experimental Medicine, by A. D. Slonim[3] and M. E. Lobashev at the Pavlov Physiological Institute, by the laboratory of A. B. Kogan (1959) at the Rostov University, and by numerous laboratories of other scientific establishments.

Subjects and Methods of Investigation

The subjects of our investigations are: fishes, tortoises, pigeons, rooks, hens, ducks, rabbits, rats, dogs, macaques (Macacus rhesus), green monkeys (Cercopitecus aetiops), baboons (Papio hamadrias), capuchins (Cebus apella), chimpanzees (Pan shimpanse), and human beings.

In choosing this series of subjects we proceeded on the basis of the gradual progressive complication of their brain structure. The degree to which the given animal could be investigated in laboratory conditions was also of importance.

[3] See the monograph of K. M. Bykov and A. D. Slonim (1960).

In natural conditions all the above mentioned animals procure their food with the help of their extremities—teeth, beaks, and jaws—they dig up the ground, turn over stones, break the branches of trees, etc.

According to comparative morphological investigations carried out by A. N. Severtsev, characteristic of all the vertebrates is an oral cavity apparatus of the grasping type. "This apparatus proved to be an extremely plastic organ, i.e., an organ capable of adapting itself to highly diverse food variations" (Severtsev, 1934, p. 29). However, already in lower vertebrates (amphibians and reptiles) the fore extremities serve as auxiliary organs for grasping food, while in higher animals (for example, in monkeys) they turn into the most important apparatus of food procurement.

In view of this great importance of the motor function of the oral cavity apparatus and of the extremities, we have applied the method of food-seeking conditioned reflexes (Baru, 1953a; Chernomordikov, 1953; Malinovsky, 1952a; Prazdnikova, 1953a; Voronin, 1954a). According to this method, a definite accidental or experimentally provoked movement performed by the animal, such as pressing a button or a pedal with one of the extremities or with the beak, catching a ring or a suspended bead with the jaws, was accompanied by the act of feeding

After the animal performed a corresponding movement at the sight of a certain manipulator (a pedal, a ring, a bead, etc.), food was administered only if the animal's action coincided with a definite external signal (auditory or photic). Thus, in response to the action of a light signal, for example, the experimental fish caught a bead with its jaws; the turtle seized a ring; the bird began to

peck carrion; the dog pressed a pedal with its forepaw; the monkey pressed a lever, a button, or some other manipulandum.

The action of the stimulus and the conditioned motor reflex were recorded on a moving band of the kymograph.

A similar motor conditioned reflex, for example, the reflex of pressing a button or a lever, was established also in man. In this case, however, it was reinforced not by food, but by the action of a signal stipulated by a preliminary instruction; the subject was told that if he performed the right movement, this would be confirmed, for example, by the emergence of a red light (Rokotova, 1954b).

Food-seeking conditioned reflexes were established not only to single sound or light stimuli, but also to complex stimuli consisting of simultaneously or successively acting sound and light agents.

Finally, with the view of studying the laws of analysis and synthesis of proprioceptive stimulations arising in the course of motor reactions, we established chains of motor conditioned reflexes, i.e., series of consecutive movements in response to chains of successive sound and light stimuli.

The formation of chains of motor conditioned reflexes was effected in two ways: either as a result of a consecutive combination of already established homogeneous or heterogeneous food-seeking movements (Rokotova, 1954a; Voronin, 1947), or by a consecutive addition of accidental or experimentally provoked movements to previously formed food-seeking reactions (Voronin & Napalkov, 1959, 1960).

In the first case, for example, the experimental rabbit or dog reacted to the light of an electric bulb by catching a ring with the teeth; at the sound of a bell it tapped a pedal with its forepaw; at the sound of a metronome it made a movement with the jaws aimed at catching another ring. When all the conditioned agents were combined into a chain of stimuli, the animal reacted to it with a corresponding series of movements, after which it was given food.

In the second case, the light of an electric bulb was switched on after the accomplishment of a previously established reaction, for example after tapping a pedal with the fore extremity in response to the sound of a bell; if the animal performed another movement which corresponded to the scheme of the experiment, this was reinforced by the administration of food. In this way a chain of two movements was formed. The same principle was applied in the establishment of a chain reaction consisting of three, four and more movements. A conditioned inhibitor may be formed to such a chain of movements by means of a systematic nonreinforcement of the chain reaction if the latter takes place in the presence of a certain (conditioned inhibitory) agent. Then it is possible to establish a chain of two movements which are provoked by external stimuli and which eliminate the conditioned inhibitory agent; after that the positive conditioned signals and the food-seeking movements become effective.

The same principle was applied in the establishment of still more complex chain reactions consisting of alimentary and defensive movements. For example, an alimentary chain of movements consisting of three components has been established in a dog; but the animal does not perform this series of movements in the presence of an acting signal which is followed by a harmful stimulus, for example, by a blow on the back with a wire clamp. By switching off the signal of the harmful stimulus and by

utilizing this as a reinforcement, we can establish a new chain of movements; the animal performs a consecutive series of movements which lead to the elimination of the signal of the harmful stimulus and, consequently, of the stimulus itself. After that the unimpeded performance of a chain of food-seeking movements by the animal leads to the obtainment of food.

In a human subject we establish in these cases not food-seeking reactions, but movements aimed at solving certain tasks stipulated by preliminary instructions. Thus, the subject was asked to switch on a definite signal (for example, a red electric bulb) by means of performing any suitable movements on the switch panel. It is this bulb that served as a reinforcement, as a signal for the correctness of the subject's movements.

If the subject performed a series of movements which corresponded to the established scheme of the experiment, these movements were reinforced. Besides, conditioned inhibitors of three kinds were established to the chain of movements; a successive elimination of each of these inhibitors with the help of two or three movements made it possible to solve the task set by the preliminary instruction.

Elementary Food-Seeking Reactions

Numerous investigations carried out in Pavlov's (1926) laboratories showed that the formation of positive conditioned reflexes (reinforced by food or by some other unconditioned stimulus) and of inhibitory or negative (unreinforced) reflexes was based on the interaction of the processes of excitation and inhibition in the brain. Both the excitatory and inhibitory processes possess definite strength and mobility, and are equilibrated. By establishing the rate of formation of positive conditioned reflexes we can determine the strength of the process of excitation and its equilibrium with the process of inhibition. The rate of formation of a negative reflex, i.e., of any kind of internal inhibition (extinction, differentiation, conditioned inhibition, retardation) may serve as an index of the strength of inhibition and of its equilibrium with the process of excitation.

In this connection, a series of investigations of elementary food-seeking conditioned reflexes in vertebrates of different levels of phylogenesis were undertaken in our laboratory. The results of numerous researches carried out by our coworkers and published in various magazines, as well as in special articles written by myself (Voronin, 1954a, 1954b, 1955) show that the rate of formation of active conditioned reflexes in vertebrates is approximately the same, and consequently cannot reflect the level of phylogenesis. The data presented in Table 1 corroborate the point of view of D. A. Biryukov (which was expressed several years prior to us) that the rate of formation of a motor reflex, such as the movement of the experimental animal towards the place of feeding at a definite signal, is not a reliable criterion for ascer-

TABLE 1

Rate of Formation of a Conditioned Reflex

Animals	Number of animals	Number of combinations after which the conditioned reflex emerged	Number of combinations after which the conditioned reflex stabilized
Goldfish	29	2–35	4–89
Hen	15	2–20	12–102
Rabbit	36	2–29	13–120
Dog	21	3–87	5–190
Green monkey, baboon, macaque	14	3–35	7–90
Chimpanzee	5	3–15	8–25

Note.—This table, as well as all others, show the limits within which the rate of formation of a conditioned reflex varies in each animal.

TABLE 2

RATE OF FORMATION OF DIFFERENT KINDS OF INTERNAL INHIBITION

Animals	Number of applications of the stimulus after which internal inhibition was formed		
	Extinction	Differentiation	Conditioned inhibition
Goldfish	28–78	20–40	15–20
Steppe and Greek tortoises	6–15	12–40	7–143
Jackdaw	—	20–40	6–8
Hen	11–84	10–30	5–15
Rabbit	10–30	6–13	12–26
Dog	10–18	4–18	—
Baboon, macaque	7–15	3–25	8–12
Chimpanzee	8–10	9–30	3–6

Note.—The double figures indicate the limits within which the rate of development of internal inhibition varies. The number of animals are the same as in Table 1.

taining the specific features of higher nervous activity of different levels of development.

A comparison of the rates of formation of negative conditioned reflexes shows that in this case a certain difference is in evidence (Table 2).

Thus, the strength of the excitatory process in vertebrates in the case of an elementary food-seeking conditioned reflex is equal. As to the strength of internal inhibition, it somewhat increases in the ascending line of vertebrates, as shown by experiments performed in similar conditions.

In view of this, one might draw the conclusion that in the course of phylogenesis the equilibration of excitation and inhibition somewhat increases. However, this conclusion is disproved by the fact that the process of excitation in monkeys considerably predominates over the process of inhibition, owing to which these animals are highly excitable and unequilibrated. Thus, we can speak only of a certain tendency towards an increase in the strength of inhibition and equilibration of the nervous processes which in a number of cases may not overstep the limits of individual variations in animals of any species,

In order to reveal certain distinctive features in the higher nervous activity of different species of animals, we also investigated the mobility of the nervous processes, the capacity of each of these processes to turn into its opposite, i.e., the capacity of excitation to turn into inhibition and vice versa.

We determined the mobility of the nervous processes by way of changing the signaling properties of the positive and negative conditioned reflexes. For this purpose, after the stabilization of the reflexes the sequence of their reinforcement was changed; the positive reflex was not reinforced at all, while the negative reflex was. Owing to this, after the lapse of a certain time the signaling properties of the stimuli changed into their opposites, and consequently, the conditioned reflexes themselves were, in our laboratory terminology, "reversed." In order to determine the capacity of the animals to endure an overstrain of the mobility of the nervous processes, and to establish its susceptibility to training, the conditioned reflexes were consecutively reversed several times in each species of animals.

Table 3 shows that in most animals, judging by the first "reversal" of the reflexes, the mobility of the nervous processes is almost equal. Only the monkeys, and particularly the chimpanzees, constituted an exception: in this case the reversal was accomplished very rapidly.

However, further changes in the signaling significance of the stimuli showed that fish and tortoises differ greatly from other animals: the re-

TABLE 3

NUMBER OF STIMULATIONS NECESSARY TO "REVERSE" CONDITIONED REFLEXES

Consecutive reversals	Fish	Tortoises	Jackdaws	Rabbits	Dogs	Baboons	Chimpanzees
1	30	25–30	41–120	47–107	33–36	9–58	4–6
2	151	21–124	26–91	17–32	2–11	2–17	3–4
3	failure	22–25	21–70	31–90	4–27	33–42	2–4
4	—	23–25	21–70	25	2–24	2	1
5	—	failure	35	13	3–19	2–3	1
6	—	—	—	11	1–2	1–2	1
7	—	—	—	49	1–2	1–2	1
8	—	—	—	—	2–4	2	—
9	—	—	—	—	2	1	—

Note.—The sign (—) used in this and other tables means that no test was performed.

flexes proved to be reversed only in one crucian out of the three which were subjected to experimentation; in the two others no reversal took place even for the first time, because after 178–200 stimulations the conditioned reflexes disappeared. In the case of tortoises only two (out of five) exhibited a reversal of the reflexes; in the other tortoises after 70–100 stimulations the conditioned reflexes began to manifest themselves very irregularly and, finally, fully disappeared.

Thus, our investigations of the mobility of the nervous processes in a series of vertebrates have revealed marked distinctions in this highly important property of the nervous processes on different phylogenetic levels.

On the basis of these facts we considered it possible to draw the following conclusion (Voronin, 1953):

> the more developed the nervous system of the animal and the greater the mobility of the processes of excitation and inhibition, the more rapidly can the animal be trained, the more resistant is the nervous system to the frequent and sharp changes of external influences (p. 53).

Our experiments have also shown that whereas in the case of higher animals the process of reversing the signaling properties of the stimuli of food-seeking reflexes (with which we were concerned in our experiments) presents a very elementary task, for lower vertebrates (crucians, tortoises) it is, apparently, a rather complex procedure which approximates the limit of their capacity to "learn" in the given conditions.

This conclusion is particularly corroborated by repeated reversals of the signaling properties of stimuli, i.e., in conditions of training, or to be more prescise, in conditions involving an overstrain of the lability of the nervous processes.

In this connection we specially investigated the process of training extinctive inhibition in fish, tortoises, hens, and rabbits (Polivannaya, 1960); in dogs and rabbits (Beznosikov, 1954); and in baboons (Voronin, 1951).

The investigations showed that is the extinction of the conditioned reflex is performed day by day, a lesser and lesser number of repeated stimulations unreinforced by food is required to extinguish the conditioned reflex. Finally (after a certain number of experimental days), it is sufficient to apply the unreinforced stimulus only once, and no subsequent application will evoke the con-

TABLE 4

TRAINING OF EXTINCTIVE INHIBITION

Animals	Number of experiments required for training
Crucian	68
Tortoise	50
Hen	36–54
Rabbit	12–38
Dog	7–8
Baboon	10–12

ditioned reflex reaction.

As shown in Table 4, in dogs and baboons this phenomenon manifests itself after a period of 7–12 days, while in lower vertebrates (crucians, tortoises) a considerably greater number of experimental days is required.

When the rate of training extinctive inhibition was established, we had to find out whether this process exerted influence on the extinction of reflexes from other analyzers. For this purpose, we compared the dynamics of extinction of reflexes to sound and light stimuli before and after training extinctive inhibition of a reflex to one of the stimuli.

The results of some of these experiments (Beznosikov, 1954; Polivannaya, 1960) are given in Table 5, which show that the training of extinctive inhibition of one reflex exerts positive influence also on the process of extinction of another reflex in animals in different levels of phylogenesis. For example, in the case of Crucian Number 5 (Table 5) the first extinction of a conditioned reflex required 37 applications of a green light, 21 applications of a red light, and 78 applications of a bell. After the reflex was extinguished chronologically, day by day, the phenomenon of extinction began to manifest itself right after the first nonreinforcement. This training of

TABLE 5

EFFECT OF TRAINING EXTINCTIVE INHIBITION OF ONE REFLEX ON THE RATE OF EXTINCTION OF OTHER REFLEXES

(Number of applications of the stimulus)

Animals	Stimulus to which the reflex was chronically extinguished	The stimulus tested																			
		Before training									After training										
		Green light	Red light	Flickering light	Bell	Revolving figure	Metronome	White light	Gurgle	Siren	Green light	Red light	Flickering light	Bell	Revolving figure	Metronome	White light	Gurgle	Siren
Crucian No. 5	Green light	37	21	—	78	—	—	—	—	—	1	6	—	9	—	—	—	—	—
Tortoise No. 10	Green light	16	19	—	—	29	—	—	—	—	1	6	—	—	2	—	—	—	—
Hen No. 63	Green light	22	22	—	59	—	—	—	—	—	1	5	—	5	—	—	—	—	—
Rabbit "Little Tom"	Green light	7	—	10	14	—	—	—	—	—	1	—	1	7	—	—	—	—	—
Dog "Gaichi"	Gurgle	—	—	—	6	—	10	8	16	—	—	—	—	2	—	4	1	1	—
Baboon "Danube"	Siren	—	1	—	5	—	—	2	—	3	—	1	—	1	—	—	1	—	1

the process of extinction of a conditioned reflex to a green light greatly hastened also the process of extinction of reflexes to other stimuli.

The influence exerted by the training of extinctive inhibition proves to be strongest in the extinction of a reflex elaborated fom one and the same analyzer.

Up to the present time the mechanism involved in the training of extinctive inhibition is not quite clear to us. We believe that the matter is reduced to the formation of a conditioned reflex to a certain attribute of the experimental procedure which signals the absence of a reinforcement. It is the first application of a conditioned reflex without any subsequent reinforcement that becomes such a signal. After repeated applications of the stimulus without reinforcement this turns into an attribute signaling the absence of food. It is probable that the role of such a signal is played by the segment of time after which the subsequent conditioned reflex begins to act; for example, in our scheme of extinction of a conditioned reflex this took place in 30 seconds. Consequently, the time factor may also become a conditioned signal.

As is known, the extinction of a reflex is more rapid the greater the lability of the nervous processes. Owing to this, experiments with training extinctive inhibition corroborate the results of our comparative-physiological investigations of the lability of nervous processes and our conclusion that the lability of excitation and inhibition perfected itself in the course of evolution. Moreover, they allow us to draw the same conclusion concerning the strength of the inhibitory process as well as a similar hypothetical conclusion concerning the excitatory process.

Besides data relating to the perfection of the excitatory and inhibitory processes in the course of evolution, we were able to accumulate some material concerning one of the most important adaptive properties of the nervous system—the stability of the stimulations that are imprinted in it, i.e., concerning the persistence of the traces of stimulations.

This property underlies both the coupling functions of the higher divisions of the central nervous system, i.e., the formation of temporary connections of conditioned reflexes, and the preservation of these reflexes, i.e., memory.

The question of the persistence of traces of stimulations first arose for us when we studied the phenomenon of conditioned inhibition in animals in different levels of phylogenesis (Chernomordikov, 1953; Firsov, 1953; Malinovsky, 1953a; Prazdnikova, 1953b). It was established that if a conditioned inhibitory agent begins to act prior to a positive stimulus and if the interval between them is equal to 5–10 seconds, no conditioned inhibition is formed in lower vertebrates (fish and tortoises). We assume that the faculty of the nervous system of these animals to preserve traces of stimulation is very little developed, so that 5–10 minutes after the exclusion of the additional agent its traces cannot acquire a conditioned-inhibitory signaling property. In order to verify this assumption, we carried out a comparative-physiological study of a conditioned reflex to time; as is known, this reflex is assigned to the group of trace conditioned reflexes, i.e., temporary connections formed to the traces of stimulation. To establish this reflex, we fed the animal at definite intervals, for example, every 2 minutes, and all the accidental or experimentally induced movements which coincided with the act of eating became conditioned.

Table 6 shows that the rate of formation of a conditioned reflex to

TABLE 6

RATE OF FORMATION AND STABILIZATION OF A CONDITIONED REFLEX TO TIME

Animals	Number of experimental animals	Number of combinations (maximum and minimum) required for:	
		First manifestation of the conditioned reflex	Stabilization of the conditioned reflex
Crucian	5	9–15	failure
Tortoise	8	24–40	failure
Hen	12	7–45	132–250
Rabbit	8	4–50	100–150
Dog	5	9–15	60–120
Monkey (macaques, baboons)	14	5–13	90–120

time, with an interval of 1 or 2 minutes, is almost the same in all of our experimental vertebrates. However, the stabilization of the conditioned reflex was dependent on the phylogenetic level of development of the nervous system (Baru, 1953; Bolotina, 1953; Chernomordikov, 1953; Prazdnikova, 1953). Thus, in crucians and tortoises the conditioned reflex did not become stabilized at all; it remained unstable and was masked by intersignal reactions; in a number of cases it was absent in spite of 200 combinations of the animal's movement with feeding at intervals of 1 or 2 minutes. With great difficulty, i.e., only after 100–250 combinations, was the conditioned reflex to time stabilized in 50% of all the experimental rabbits and hens.

The conditioned reflex was most rapidly stabilized in dogs and monkeys. Besides, in these animals, in comparison with others, the reflex was more susceptible to external inhibition and its extinction was more protracted. Finally, the duration of the time interval to which the conditioned reflex could be formed best proved to be different in various animals. For example, in fish, tortoises, hens, and rabbits the most distinct reflexes were formed to intervals of 1 or 2 minutes; in monkeys—to intervals of 2–5 minutes; and in dogs—to intervals of 4–6 minutes. In monkeys and dogs a conditioned reflex could be established to an interval of 11 minutes, while in our other experimental animals the number of adequate reactions in the course of the experiment already showed a decline at an interval of more than 2–5 minutes, and subsequently fully disappeared.

Thus, we observed a direct dependence of the rate of stabilization of a conditioned reflex to time, as well as the degree of its distinctness and the conditions of its formation, on the phylogenetic level of development of the nervous system; this allows us to draw the following conclusion: the more developed the nervous system, the longer it retains traces of stimulation.

This point of view is corroborated, along with the above mentioned facts, by the results of accidental and other observations in our laboratory on the persistence of conditioned reflexes after certain intervals in their training.

For example, we observed a complete recovery of a complex system of motor conditioned reflexes in a rhesus macaque after an interval of 8 years in experimentation (Voronin & Shirkova, 1949).

The persistence of food-seeking conditioned reflexes was observed in hens after an interval of 3 months; fish manifested a deranged differentiation of inhibition after an interval of 3–5 days (Prazdnikova, 1953a).

Special experiments performed on crucians, pigeons, and rabbits (Chumak, 1958) showed that the higher the phylogenetic level of development of the nervous system, the less influence is exerted by an interval in experimentation on the preservation of a food-seeking conditioned reflex.

Thus, on the basis of the experi-

mental material we can state that the food-seeking elementary conditioned reactions, with which we dealt in our experiments, constitute a universal phenomenon in vertebrates, a phenomenon that is formed in the individual life of the animals according to the laws of the conditioned reflex. This phenomenon is so universal and at the same time so elementary that in this respect it is impossible to establish any distinctions with respect to different levels of phylogenesis of the higher nervous activity of vertebrates.

Apparently, the properties of the excitatory and inhibitory nervous processes on all the phylogenetical levels that were investigated by us are so developed that they insure the accomplishment of these elementary food-seeking reactions—reactions which adapt the organism to its usual conditions of existence. But, as shown by special experiments, any sharp change in these conditions exerts different influences on the conditioned reflex activity of various animals. For example, overstrain of the strength and lability of nervous processes makes it possible to reveal marked phylogenetical distinctions.

The fact is that the more these properties of nervous processes are developed, the more developed is the higher nervous activity of the animals. This suggests the idea that it is the perfection of the properties of the excitatory and inhibitory nervous processes that lies at the base of the evolutionary perfection of higher nervous activity. In its turn, the development of the properties of nervous processes underlies the perfection of the general mechanisms of higher nervous activity. We endeavored to disclose this interrelation by means of a comparative-physiological study of different kinds of conditioned reflexes. Thus, we investigated conditioned reflexes to time (Bolotina, 1952a, 1952b, 1953; Voronin, 1951; & others), temporary connections between indifferent stimuli (Malinovsky, 1953b; Rokotova, 1952, 1953, 1954a), conditioned reflexes of second order (Brau, 1953b; Malinovsky, 1952b; Prazdnikova, 1953), imitative conditioned reflexes (Bogomolova, Saakyan, & Kozorovitsky, 1956; Kozorovitsky, 1956; Voronin, 1947b), and conditioned reflexes to relative attributes of stimuli (Chumak, 1957a, 1957b).

In the course of analysis of the results of these investigations we were

TABLE 7

EFFECT OF INTERVALS IN EXPERIMENTATION ON THE PRESERVATION OF FOOD-SEEKING CONDITIONED REFLEXES

Animals	Number of combinations performed before the interval in experimentation		Number of combinations required for recovery of the conditioned reflex to its initial level after the interval	
	10 days	20 days	10 days	20 days
Crucian No. 6	900	1050	20	71
Crucian No. 7	780	1250	35	84
Crucian No. 8	780	1250	43	74
Pigeon No. 1	220	270	5	22
Pigeon No. 3	—	460	—	23
Pigeon No. 7	555	610	7	23
Rabbit No. 1	890	600	0	5
Rabbit No. 4	840	610		5

primarily impressed with the universal significance of the food-seeking reflex, irrespective of its kind. However, separate facts obtained as a result of our investigations showed that the principal mechanism in higher nervous activity which developed during the course of phylogenesis was the mechanism of analysis and synthesis of stimuli, as well as that involved in the activity of the organism arising in response to stimuli.

Analysis and Synthesis of External Stimuli

An organism at any level of phylogenesis is subjected to a tremendous number of influences exerted by the external environment; but not all organisms possess equally the faculty to isolate and analyze the elements of these influences, to synthesize them into complexes, and to link them with one of their own activities.

In other words, not all organisms are equally capable of learning to adapt themselves to the external environment in the best possible way on the basis of the complex signals that come from this environment.

As far back as the early years of research (1903–1910) Pavlov's laboratories demonstrated that the salivary reflex is formed to a complex of stimuli, that the animal is able to distinguish a separate component from the whole complex and that after long training the complex is perceived as a single whole (Pavlov, 1926). These investigations laid the foundation of further thorough research into the analytico-synthetical mechanisms of higher nervous activity. From the point of view of the above, we were interested to find out the degree of development of these mechanisms in vertebrates at different levels of phylogenesis. With this end in view, and using the same index of conditioned reflex activity, namely, the food-seeking reaction, we began to apply as conditioned stimuli various combinations of agents addressed to the same or different analyzers, i.e., to the sense organs in the broad meaning of the word, including their cortical projections. Depending on the tasks in our investigation, we communicated either an active or an inhibitory signaling property to combinations of stimuli consisting of simultaneously or consecutively acting agents.

Several variants of conditioned reflexes to complex stimuli were investigated in our laboratory in the course of a number of years. In particular, our laboratory investigated a simple case of a complex inhibitory conditioned stimulus which is termed "conditioned inhibitory combination of stimuli" (Chernomordikov, 1953; Firsov, 1953; Malinovsky, 1953a; Prazdnikova, 1953b).

In some cases the conditioned inhibitory agent was added to the positive stimulus and they acted simultaneously (a simultaneous combination). In other cases the conditioned inhibitory agent acted first; then its action was discontinued and a positive agent took its place (a consecutive combination). Finally, in a third series of experiments these two stimuli were separated from each other by an interval of 5–20 seconds (a consecutive-trace combination). In all these cases the combinations of stimuli were not accompanied by food, as is usually done in the case of formation of a conditioned inhibitor.

Table 8 shows that the possibility and rate of formation of a conditioned inhibitor are not the same in different animals and in the presence of different conditioned inhibitory combinations of stimuli.

Whereas under the simultaneous action of a combination of stimuli conditioned inhibition is formed in all animals, in the case of a consecutive

TABLE 8

RATE OF FORMATION OF DIFFERENT VARIANTS OF CONDITIONED INHIBITION

Animals	Number of applications of the conditioned-inhibitory combination			
	Simultaneous combination	Consecutive combination	Consecutive-trace combination of stimuli	
			Interval between stimuli	
			5 seconds	10 seconds
Goldfish	15–20	60	69	—[a]
Tortoise	7–143	—[a]	—[a]	—[a]
Hen	5–15	10–70	99	68–110
Rabbit	14–26	40	58	69–80
Baboon	8–11	8–120	20–100	35–41
Chimpanzee	3	8	2	2

[a] Conditioned inhibition was not formed.

action of stimuli this kind of internal inhibition develops in lower vertebrates (fish and tortoises) with great difficulty, or does not emerge at all. Only higher apes constitute an exception: in this case a negative temporary connection is invariably formed under any of the tested structures of complex inhibitory stimuli.

The above facts show that the analysis and synthesis of a complex inhibitory stimulus is feebly pronounced in fish and tortoises. These processes are somewhat better accomplished in hens and rabbits, still better in baboons, and they are greatly developed in higher apes—chimpanzees. Then, with the same aim in view, we investigated a more complicated form of differentiation of complex stimuli, in which one and the same agent acquired either a positive or a negative signaling property, depending on the stimulus with which it acted simultaneously. For example, a positive conditioned reflex was established in the experimental animal to Agent A and a negative conditioned reflex to Agent B. Then an indifferent Stimulus C was added in some cases to Stimulus A and in others to Stimulus B. The combination of Stimuli A+C was not reinforced by food, while the combination B+C was. After a number of applications the combination A+C ceased to evoke a conditioned reaction, i.e., C began to play the role of a conditioned inhibitor which inhibited the positive signaling property of A. At the same time the combination B+C began to evoke a conditioned reflex, i.e., C acquired the significance of a conditional disinhibitor which eliminated the inhibitory property of B.

Our collaborators, S. M. Makokina and Y. A. Kholodov (1959), investigated this phenomenon with chimpanzees, baboons, and dogs; and K. I. Iordanis (1959a) investigated it with rabbits, pigeons and tortoises. It was found that a differentiation of the double property of the same agent could be established in chimpanzees, baboons, and dogs, though with great difficulty. Owing to this, the indifferent agent which was added to the positive conditioned stimulus began to inhibit its effect, while the indifferent agent added to the negative stimulus began to eliminate its inhibitory action.

This differentiation is less pronounced in rabbits and pigeons; we have never seen a single animal of this group properly respond to a

conditioned-inhibitory and a conditioned-disinhibitory combination of stimuli within the same experiment and in all, or almost all, cases. Both "proper" and "improper" reactions were usually observed in the course of the same experiment.

Experiments with Greek tortoises did not produce even a hint of any differentiation by these animals of the complex stimuli that were applied by us.

Thus, in this case of differentiation of complex conditioned reflexes, just as in the previous case, we observed a similar picture of marked distinctions in the degree of development of the analytico-synthetical mechanisms of lower and higher vertebrates. This conclusion is also corroborated by data obtained from investigations on conditioned reflexes to simultaneous complexes of stimuli and to chains of stimuli.

It was long ago established by the investigations in Pavlov's laboratories, as well as by other investigations, that if a complex conditioned reflex is established in a dog to a combination of stimuli, for example, to three simultaneously acting stimuli, at first both the whole complex of stimuli and its separate components will have signal significance. But with the training of the conditioned reflex, separate components gradually lose their signaling properties, and only the complex of stimuli as a whole can evoke a conditioned reflex effect.

Investigations carried out in our laboratory (Voronin, 1957) on monkeys (baboons), rabbits, and fish showed that this process of synthesizing separate components of a complex of stimuli into a single, integral stimulus is inherent in monkeys, but is wholly absent in fish.

These investigations disclosed that a conditioned reflex to the simultaneous action of two (sound and light) stimuli is established in monkeys and dogs quite rapidly after 3–20 combinations. In monkeys it is also easy to establish a differentiation between a complex of stimuli and its separate components. For this purpose, it is sufficient to use seven isolated applications of the components without food reinforcement.

The establishment of such a differentiation is much more difficult when separate components—and not the complex as a whole—possess a positive signaling property. This requires about 100 contrapositions of the complex of stimuli unreinforced by food and of its components which are accompanied by the administration of food.

Whereas in monkeys both variants of a differentiation between a complex of stimuli and its components could be established with greater or lesser difficulty, in fish they are not observed at all.

Rabbits occupy an intermediate place between monkeys and fish; although it proved possible to establish in them with great difficulty a differentiation between a complex of stimuli applied without food reinforcement and its components accompanied by the presentation of food, this differentiation was very unstable; the animal often reacted to the complex of stimuli in a positive way.

Motor conditioned reflexes to chains of stimuli were investigated in fish (Prazdnikova, 1955), tortoises (Voronin, 1957), hens, pigeons, jackdaws, and crows (Ovchinnikova, 1955), dogs (Firsov, 1954), chimpanzees, macaques, baboons, and capuchins (Firsov, 1955; Prazdnikova & Firsov, 1953; Shirkova, 1949).

The scheme of a three-component chain of stimuli consisted in the following: the first component acted during 5 seconds, then it was excluded and replaced by the second

component which acted during the succeeding 5 seconds; after that the third component of the chain stimulus which coincided with the administration of food was put into action for a period of 10–15 seconds. This scheme was applied in all experiments, but the time of action of the stimuli was in some cases prolonged, especially when slow animals, for example, tortoises, were investigated.

In studying these conditioned reflexes we paid special attention to the phenomenon which has been mentioned above in connection with experiments concerning conditioned reflexes to a simultaneous complex of stimuli, namely, to the process of synthesis of separate components into a single stimulus.

In order to reveal this process, we used to test from time to time the signaling properties of separate links of the chain of stimuli in the course of protracted training of conditioned reflexes in various animals.

As seen from Table 9, in fish, in spite of frequently observed negative reactions to isolatedly applied components of chains of stimuli, in most cases they do not lose their signaling properties, even after a very long training of a conditioned reflex to a complex stimulus.

On the basis of these data we can state that crucians manifest a certain feebly pronounced tendency to differentiate a chain stimulus from its components. In tortoises such a tendency is out of the question, since only two cases (out of 30 tests performed) produced a negative effect, which may be accounted for by accidental circumstances.

Mammals (beginning with rabbits) differ greatly in this respect from fish and tortoises. Table 10 shows that in the overwhelming majority of cases the first and second links of the chain of stimuli lost their signaling property for these animals, and only the third link retained it longer than the two others. Experiments on dogs yielded almost similar results,

A comparison of the results obtained from experiments on rabbits and dogs with those obtained from investigations of the lower and higher monkeys revealed a striking difference between these groups of animals.

Experiments with lapunder macaques showed a very rapid (in the course of only 3 experiments) development of the retardation of a reflex

TABLE 9

TESTS OF THE SIGNALING PROPERTIES OF SEPARATE COMPONENTS OF A CHAIN STIMULUS IN FISH
(Experiment with two crucians)

Components	Tests						
	1	2	3	4	5	6	7
I	101+	193−	306+	516+	604+	783−	1102+
	101+	194−	310+	437	631−	1079+	1339+
II	104+	196+	309+	523−	586+	1080−	1104−
	104+	197−	317+	439−	634−	792−	1343+
III	105+	199−	312+	534+	583+	789−	1106+
	107+	199+	320−	441−	636−	1083+	1347+

Note.—The figures designate the number of applications of the chain stimulus after which tests were performed; the + sign denotes the presence of a conditioned reaction, the − sign denotes the absence of a reaction.

TABLE 10

TESTS OF SIGNALING PROPERTIES OF SEPARATE COMPONENTS OF A CHAIN STIMULUS IN RABBITS

Number of rabbit	Components		
	I	II	III
1	66+	70−	74+
	687−	690−	693−
	775−	776+	781−
	1204−	1202+	1209−
		1210+	1212−
3	65−	67−	69+
24	83−	80−	78+
	118+	120−	123+
	173−	174−	175+
25	78+	79+	80+
	109+	112+	118+
	185−	186−	187+
	221−	219+	223+
	336−	338+	340−
26	177−	175−	176+

Note.—Plus and minus signs have the same meaning as in Table 9.

to the first components of the chain of stimuli. Tests of separate components revealed the absence of a conditioned reaction after 40–60 applications of the chain stimuli. The results of these experiments were also corroborated in experiments with capuchins, although these monkeys, which are more excitable than macaques, in most cases exhibited a positive reaction if the components acted during the same period as the chains of stimuli.

Thus, the results of experiments on fish, tortoises, rabbits, dogs, and monkeys led us to the conclusion that the more developed the nervous system of the animal, the better is the synthesis of the components of a complex stimulus effected. This conclusion is based on a comparison of data obtained from experiments with conditioned reflexes to different kinds of complex stimuli, including chains of stimuli. It should be borne in mind, however, that the process of synthesizing separate links of a chain of stimuli into a single stimulus does not always manifest itself quite clearly judging by the fact that separate components lose their signaling properties. This is apparently due to the specific nature of the synthesis of a chain stimulus: in this case each of the components first of all directly signals the appearance of the next component and indirectly signals the appearance of food. Owing to this, only higher animals can perceive an isolated application of a component of a chain of stimuli as a new stimulus.

This process of differentiation is accomplished through the mechanism of the orienting reflex which is evoked by the difference between the action of the component and the usual action of the complex of stimuli.

Comparative-physiological investigation of conditioned reflexes to different kinds of complex stimuli is of double significance. First, it shows that the fusion of separate influences falling upon the organism into a single complex stimulus is effected gradually, according to the same principle of analysis and synthesis which underlies the formation of an elementary conditioned connection. Second, it allows us to make the assumption that the faculty of analyzing and synthesizing stimulations has developed in the course of evolution.

It is clear that the material accumulated by us during the study of conditioned reflexes to complex stimuli does not give any ground for concluding that lower vertebrates do not possess the faculty of analyzing and synthesizing complex stimuli altogether. The fact that conditioned reflexes to complex stimuli were established in such animals

testifies to the analytico-synthetical activity of the higher division of their nervous system. But this activity has not yet assumed the higher form due to which numerous and complex interacting "functional combination centers" (Pavlov, 1938, p. 365) are formed in the brain structures where the function of coupling is effected.

Owing to this faculty of higher adaptation, the organism is able, as it were, to "group" or "reduce" the infinite number of stimulations that fall upon it and to associate them, according to the principle of temporary connection, with its own various activities.

Analysis and Synthesis of Proprioceptive (Kinesthetic) Stimulations

When we deal with such a complex and highly universal external manifestation of higher nervous activity as motor activity acquired by the organism in the course of its individual life, we must possess definite knowledge of the analysis and synthesis not only of exteroceptive, but also propioceptive stimulations.

On the basis of a series of investigations conducted in Pavlov's laboratory (1938) it has been assumed (in general form this was expressed by I. M. Sechenov as far back as 1861–66) that any complex movement "learnt" by a human being or animal is none other than a chain of motor reflexes, where the end of one reflex serves as a stimulus of another reflex. In other words, the formation of a chain action in the shape of a complex motor act is based on the analysis and synthesis of proprioceptive (kinesthetic) stimulations.

Kinesthetic stimulations interact with the exteroceptive ones (auditory, photic, and others) and can, of course, associate with interoceptive signals. Apparently, this explains the fact that the formation of two homogeneous[4] food-seeking reflexes to the light of electric bulbs of different colors applied in one and the same spot of the experimental chamber is equally greatly impeded in carps, piegons, rabbits, and dogs (Krushinskaya, Kholodov, Shuranova, & Shcherbina, 1960). Even after 300–700 trainings of two motor reflexes, the conditioned reflexes are far from being appropriate in all cases, i.e., they do not always correspond to the stimuli. For example, in dogs and rabbits the number of reactions to both manipulators increased due to the difficulty of differentiation; in fish there was an increase in the number of reactions to one of the two manipulators, i.e., there was observed a predominance of reactions either to the manipulator situated in the left corner of the chamber, or to that situated in the right corner.

In pigeons, as well as in fish, this phenomenon of predominance was likewise observed, but it was of a less pronounced character.

As shown by the experiments of of V. I. Ivanova (1960), a differentiation of two homogeneous motor reflexes is established with comparative ease in fish, pigeons, and rabbits if the conditioned stimulus is applied in different spots of the chamber; after 50–200 applications of both stimuli all these animals exhibited proper reactions. It was found that the individual specific features of the animals manifested themselves with greater force than did their phylogenetic distinctions.

In the course of formation of three motor reflexes essential differences

[4] We term homogeneous conditioned reflex movements those which are effected by the same groups of muscles and heterogeneous those which are effected by different groups of muscles.

were observed between fish, on the one hand, and pigeons and rabbits, on the other.

In fish no complete differentiation of three reflexes was noted, in spite of 500–800 applications of each of the stimuli; erroneous reactions took place in 30–40% of all cases, and only in separate experiments did the animal properly react to all three stimuli.

Pigeons and rabbits proper reactions took place after 100–300 applications of the stimuli and in only 4–10% of all cases were erroneous movements observed. Absolute differentiations of four reflexes were established in pigeons and rabbits with great difficulty only after 400–600 applications of the stimuli. The dynamics of formation of several homogeneous motor reflexes in fish, piegons, and tortoises showed: (*a*) that the establishment of three reflexes[5] constitutes the limit of the analytico-synthetic abilities of fish, while in pigeons and rabbits four homogeneous motor reflexes are easily established; (*b*) that the formation of two and more reflexes passes through two stages: the stage of generalization, when one stimulus can evoke any of the established reflexes; and the stage of specialization, when the movement is adjusted to a strictly definite stimulus.

We also investigated the dynamics of formation of several heterogeneous motor reflexes, i.e., movements which differ from one another by the fact that they are accomplished either by different effectors—oral cavity or extremities, or by different muscles of the same effector—for example, when the subject presses a lever towards himself or in the opposite direction, or by a different number of muscles—for example, a local grasping movement by the jaws and the movement of the animal to a definite place in the chamber (Iordanis, 1958, 1959b; Voronin & Iordanis, 1960, 1961; & others).

As may be seen from Table 11, the principal difference between tortoises and other animals consists in the fact that in tortoises, just as in fish, a strongly pronounced differentiation of motor reflexes is observed only in the case of two such reflexes. In the case of three reflexes there is evidence of a disturbance of the conditioned reflex activity which is expressed in the disappearance of all previously established reflexes.

TABLE 11

COMPARATIVE DATA CONCERNING DIFFERENTIATION OF MOTOR REFLEXES

Animals	Number of experiments required for differentiating motor conditioned reflexes[a]		
	Two reflexes	Three reflexes	Four reflexes
Pigeon	3–16	6–9	—
Rabbit	10–17	12–57	41–44
Dog	18–46	8–24	—
Macaque (rhesus)	25–36	20–114	90–144
Chimpanzee	28–41	57–92	8–14

[a] In the course of experimentation 12–18 stimulations were applied to tortoises, 15–20 to pigeons and rabbits, and 30–35 to dogs and monkeys. The number of applications of each stimulus was approximately equal.

On the basis of these experiments, we can draw the paradoxical conclusion that two or three reflexes are better differentiated in pigeons than in rabbits, while in monkeys they are differentiated worse than in rabbits. A differentiation of four motor reflexes proceeds with greater difficulty in monkeys. Chimpanzees constitute

[5] Experiments performed by K. A. Iordanis showed that the possibilities of tortoises are still more limited: only two homogeneous motor reflexes could be established in this case.

a certain exception in this respect: a differentiation of four reflexes is established in them easier than a differentiation of two or three reflexes. We find it difficult to explain this difference between pigeons and rabbits. As to the difference between monkeys and other animals, they can be explained, apparently, by the fact that these animals are unequilibrated. The process of excitation considerably predominates in them over the process of inhibition (Voronin, 1952); as is known, this circumstance greatly impedes the process of differentiation of any stimuli, including kinesthetic ones.

A differentiation of several motor conditioned reflexes is established much easier in the course of one experiment if each of these reflexes, independently of all others, was previously established and stabilized in various experiments.

The investigations of Tagiev (1958) showed that if two conditioned reflexes, previously established in separate experiments, are "brought together," the conditioned-reflex activity of the animals proves to be deranged only for a short time as manifested in a larger number of erroneous reactions. But even here a certain difference was observed between pigeons and rabbits, on the one hand, and carp, on the other.

Rabbits and pigeons exhibited proper reactions after only two or three applications of the stimuli, while carp showed such reactions only after 5–60 applications.

A similar difference between these animals was clearly noted in the experiments of Chumak (1957b) who studied conditioned reflexes in relation to the magnitude of stimulation. In these experiments, the animal had to perform a certain action with two manipulators near the larger positive conditioned stimulus (reinforced by food). In fish a differentiation between the signals and the direction of movement was first observed after 110–118 applications of the stimuli; it became stabilized after 200–210 such applications. In pigeons the differentiation was first observed after 40–80 applications, in rabbits after 60–80 applications; the differentiation was stabilized in pigeons after 6–90 and in rabbits after 90–110 applications of the stimuli.

Thus, facts obtained from the study of the dynamics of establishment and stabilization of several food-seeking heterogeneous and homogeneous conditioned reflexes show that this is a more complex process than the establishment of a single food-seeking reflex.

A comparison of these facts with the previously described results of investigations on the properties of nervous processes allows us to assume that the phylogenetic difference in the analysis and synthesis of stimulations is, apparently, due to the different strength of the nervous processes of excitation and inhibition, as well as to their lability and equilibrium.

We believe that the interaction of excitation and inhibition lies at the base of any differentiation. For example, two or more manipulanda are placed in front of the animal, but only one signal emerges; in this case the animal must choose one of the manipulanda and perform a definite movement; in other words it must differentiate these manipulanda, and this, in the final analysis, means differentiation of movements on the basis of kinesthetic signals.[6] It is quite obvious that when the animal

[6] It goes without saying, that a definite role is played here also by stimulation of the receptors of other sense organs, but, as we shall see later, in the case of formation of motor conditioned reflexes kinesthetic signals predominate.

effects one movement or one motor conditioned reflex, it inhibits another movement, another motor conditioned reflex.

The stronger the nervous processes of excitation and inhibition, the more labile and more equilibrated they are, the easier is effected this coordination of movements. From this point of view, it becomes clear why fish, which possess weak nervous processes, and lower monkeys, in which the nervous processes are not sufficiently equilibrated, coped with the experimental task with greater difficulty than rabbits and pigeons. It should be pointed out that when we speak of weakness, insufficient equilibrium and poor lability of nervous processes, we apply these notions relatively to the conditions of the given experiments and in comparison with a higher level of development of the nervous system.

We can rightly say that in any species of animals, which live in adequate conditions, the properties of the nervous processes are sufficiently developed to insure their adaptation to the conditions of existence through individual training, i.e., through temporary connections or conditioned reflexes.

It must be also borne in mind that although the differentiation of stimulation is feasible due to the properties of both excitatory and inhibitory nervous processes, it is the absolute and relative strength of inhibition that plays the decisive role. Owing to this factor, fish which possess weak internal inhibition, and lower monkeys in which the inhibitory process is comparatively weak, i.e., animals with unequilibrated nervous processes where excitation predominates over inhibition, solved the experimental task with greater difficulty than pigeons and rabbits.

At the same time monkeys had an obvious advantage over lower vertebrates in another respect: while in fish and tortoises it proved possible to establish only two separate motor reflexes, in monkeys the number of such reflexes reached four, and even this was not the limit of their capacity to differentiate proprioceptive stimulations.

Up to now we have dealt with the predominant manifestation of the process of analysis, with the separation of one stimulus from two or several stimuli. This process, however, is indissolubly bound up with the process of synthesis or the process of combining stimuli.

We judged the process of synthesis kinesthetic stimulations by the dynamics of the fusion of separate motor acts into a single chain of movements. After the establishment of two or several stable conditioned reflexes we formed a chain of their corresponding stimuli in which the agents succeeded one another, each acting during 2 or several seconds.

First of all, it proved that in response to a chain of stimuli the experimental dog performed a chain of corresponding movements without the reinforcement of each of them by food (Rokotova, 1955; Voronin, 1947a).

Similar phenomena were observed in experiments with pigeons, rabbits, and monkeys in which four conditioned stimuli were combined in a chain (Iordanis, 1958; Ivanova, 1960; & others).

It was also found that in fish this process of "mechanical" synthesis[7] of stimuli and movements is quite different. When the conditioned stimulus of a separate movement was not single, but bore a chain character, the

[7] By this term we designate the phenomenon which arises without any special elaboration of the process of combining separate movements into a chain; although the animal did not undergo any preliminary training, it reacts to the conditioned stimuli in the same sequence in which they are applied.

consecutive combination of such "simple" chains of stimuli into a complex chain did not evoke any chain of corresponding movements; the fish remained beside the bead which was linked with the first chain stimulus; it performed several movements in spite of the fact that the second link of the complex chain of stimuli was already in action; then the fish swam away (Tagiev, 1958).

When a single agent was used as a conditioned stimulus of a separate reflex, the process of mechanical synthesis of movements proved feasible; however, the number of the reflexes combined was limited.

Thus, experiments performed by Ivanova (1960) on carp showed that the combination of three conditioned stimuli into a chain evoked a chain of movements of the same sequence.

This phenomenon was not noted in tortoises. To a chain consisting only of two stimuli they reacted with repeated movements corresponding to the first component of the chain; then this reaction was fully discontinued.

In experiments with rabbits and pigeons, Tagiev (1958) observed a proper reaction to a complex chain of stimuli which consisted of seven applications of two alternating simple chains. To each application of a simple chain the animal reacted with a corresponding movement.

The difference between the analytico-synthetical abilities of fish and tortoises, on the one hand, and pigeons, rabbits, dogs, and monkeys, on the other, was still more obvious when we began to apply regularly a chain of stimuli consisting of three and more sound and light agents.

In a series of experiments performed by Ivanova (1960) the training of a three-component chain of movements led to a marked disturbance of conditioned reflexes in fish; in response to a chain stimulus the animal in most cases reacted with disorderly movements, and then dropped to the bottom of the aquarium. If the fish did manifest a positive reaction to the stimuli, this reaction was of an unusual character: the fish rapidly approached the bead, snatched it violently and impetuously, and then as rapidly swam away, whereas usually the fish snatches the bead and turns to the place where the food appears (if for some reason the food does not appear, it snatches the bead again). The replacement of a three-component chain of stimuli by a two-component chain led to the normalization of the conditioned reflex activity: the fish began to manifest proper reactions to the stimuli.

After 50–70 applications of a chain stimulus there were observed some cases when the fish reacted with a two-component chain of movements just to the first component of the stimulus alone. Some of the intersignal reactions[8] also bear the character of chain reflexes. Thus, it is possible to establish in fish a system or stereotype consisting of two movements arising in a definite sequence to the first link of a two-component chain of stimuli. This proved to be the limit of the analytico-synthetical abilities of carp.

It must be pointed out that the above-mentioned experiments of Tagiev on fish demonstrated not only the absence of any so-called mechanical synthesis of movements after a consecutive application of two chain stimuli, but also the absence of real

[8] Reactions "spontaneously" arising in the intervals between conditioned stimulations. Apparently, certain elements of the surroundings in which the experiment takes place, as well as some internal stimulation (for example, "hungry" blood) are the stimuli of such reactions.

synthesis[9] after a chronic application of these stimuli in a chain order. An attempt to establish a set of two movements evoked only by the first component of a chain stimulus resulted in a disturbance of the conditioned reflex activity; the fish swam away from the stimuli.

At the time of such negative reactions to a complex chain of stimuli there was also observed a disturbance of reflexes to its components—to the simple chain stimuli: the fish did not always react to these stimuli or "confused" them, producing inadequate reactions.

These distinctions, as well as those which were revealed by the experiments of Ivanova and Tagiev in the mechanical synthesis of movements, are, apparently, accounted for by the different complexity of the conditioned stimuli.

[9] By this term we designate the phenomenon of formation of a system or stereotype of movements which arises only to the first or "starting" component of a chain stimulus.

In pigeons, rabbits, dogs, and monkeys the dynamics of formation and the length of the system of reflexes consisting of several heterogeneous motor reactions depends both on the individual properties of the nervous system and, especially, on its phylogenetic peculiarities.

Our data (Voronin & Iordanis, 1960) which were verified and supplemented in subsequent experiments (they are presented in Table 12) show, that in spite of considerable individual distinctions, the formation of a motor set of three or four movements is the limit of the analytico-synthetical abilities of pigeons and rabbits. In monkeys and apes (higher and lower), judging by the dynamics of the process of formation of an automatized chain motor reflex (a motor stereotype), the ability to synthesize four movements is not the limit.

A similar picture of synthesis of four homogeneous motor reactions into a chain of movements was ob-

TABLE 12

RATE OF SYNTHESIS OF TWO, THREE, AND FOUR MOVEMENTS

Animals	Number of stimulations required for:					
	Synthesis of two movements		Synthesis of three movements		Synthesis of four movements	
	First manifestation	Stabilization	First manifestation	Stabilization	First manifestation	Stabilization
Pigeon No. 11	15	18	52[a]	—	—	—
Pigeon No. 12	79	82	94	128	334[a]	—
Rabbit No. 21	63	95	304[a]	—	—	—
Rabbit No. 22	84	96	85	225	—	—
Rabbit No. 23	26	58	89[a]	—	—	—
Rabbit No. 25	69	unstable	169[a]	—	—	unstable
Rabbit No. 29	53	73	280	283	411	—
Dog "Marsik"	71	108	124	125	—	—
Macaque "Pashka"	67	89	88	114	124	131
Capuchin "Shalun"	25	90	91	91	94	94
Chimpanzee "Lada"	42	42	49	80	81	105
Chimpanzee "Sultan"	27	59	62	66	105	109

[a] Failure, experiments discontinued.

served by Ivanova (1960) in experiments on pigeons and rabbits. This investigation showed that a firm connection was established between the first and second, as well as between the third and fourth components of the movements. A less stable and durable connection, which is, therefore, not always manifest, is established between the second and third links of the chain of movements.

Owing to this distribution of connections within the structure of a chain motor reflex, a chain of movements is formed best to a consecutive application of the first and second components of the chain stimulus, without the application of the third and fourth. The application of the first component alone seldom evoked all four movements, as was noted in the case of synthesis of heterogeneous movements, but the second movement always emerged without the application of a corresponding stimulus.

The number of stimulations used in formation of an automatized chain of homogeneous reflexes was approximately the same as in the case of heterogeneous movements.

Synthesis of four movements proved to be of less pronounced character; this is, apparently, explained by monotonous kinesthetic stimulations which vary less than in the case of heterogeneous movements. Owing to this, the exteroceptive conditioned stimulations assumed a greater signal significance; therefore, when the third and fourth components of the chain stimulus were excluded, the chain of movements could still be effected, but when the second component was also excluded, the animal experienced certain difficulties. In the case of heterogeneous movements, i.e., when the kinesthetic stimulations differ greatly, the exteroceptive stimuli are of lesser significance; therefore after long training of the chain of reflexes, the latter is effected without any difficulty in response only to the first component of the chain stimulus.

The predominant significance (under the conditions of our experiment) of kinesthetic stimulations, as compared to exteroceptive stimulation, clearly manifested itself in the effect of an inhibitory (differential) stimulus which was introduced in different points of the chain of stimuli before and after the chain of movements turned into a system, or became automatized.

As shown by Table 13, the differential stimulus (a red light) included in the chain of stimuli prior to the

TABLE 13

DISINHIBITION OF A DIFFERENTIAL STIMULUS DEPENDING ON THE SYNTHESIS OF FOUR REFLEXES

Sequence of application of conditioned stimuli	Cases of manifestation of the inhibitory effect in percents			
	Prior to the synthesis of the chain of reflexes	Two reflexes	Three reflexes	Four reflexes
Bell, red light (differential), white light, blue light, gurgle	64	58	28	12
Bell, white light, red light (differential), blue light, gurgle	84	50	0	0

synthesis of the motor chain reaction retained its significance in 64–84% of all the tests performed. After the synthesis of four movements into a single chain reaction which is effected "automatically" only in response to the first component of the chain stimulus, the differential stimulus evoked a positive effect in all or almost all of its action (Voronin, 1960).

The phenomenon is the more pronounced, the closer the differential stimulus in the chain to the moment of food reinforcement.

From all that has been said above concerning the synthesis of motor acts into a chain reaction it follows that the longer the training of the chain of movements and the closer the component of a chain stimulus to the moment of food reinforcement, the lesser is the role of exteroceptive stimulation and the greater is the role of kinesthetic stimulation in the formation of a chain motorreaction.

This conclusion is of great importance for the comprehension of the mechanism which controls the automatization of man's motor habits—their accomplishment, as it were—without the participation of any external signals or consciousness.

The question will be considered in greater detail in a special section of this report following the discussion of facts dealing with the synthesis of more complex systems of motor conditioned reflexes established by the method of gradual complication. First of all we formed chain motor reflexes by the above described method of gradually adding new stimuli and movements to an already established reflex.

It was found that the length of the chain of movements and the rate at which a new link is added to it depend on the phylogenetic level of the animal's development. For example, an unstable three-component chain of movements could be established in tortoises with great difficulty, while in pigeons, rabbits, and cats it proved possible to establish chain reactions consisting of 7–9 components (Napalkov, 1958a, 1958b).

Table 14, based on the results of experiments conducted by a number of our collaborators (Napalkov, 1958a; Napalkov & Verevkina, 1959; Shirkova & Verevkina, 1960; & others), shows that in tortoises the formation of a new link in a chain of movements requires a considerably larger number of combinations of the new signal and of the new movement

TABLE 14

RATE OF FORMATION OF COMPONENTS IN A CHAIN MOTOR REFLEX

Animals	Number of combinations required for the formation of conditioned reflexes (average figures)				
	Single reflex	Two-component reflex	Three-component reflex	Four-component reflex	Five-component reflex
Tortoise	8	58	116	failure	
Pigeon	10	16	15	24	30
Rabbit	15	18	14	18	21
Rat	9	12	13	28	18
Cat	21	15	28	24	18
Dog	11	12	15	12	14
Baboon and macaque	2	2	2	failure	failure
Chimpanzee	2	2	2	failure	failure
Man	2	1	1	1	1

with the previously established reflexes than the establishment of any preceding link. In pigeons this phenomenon is less pronounced than in tortoises; rats and rabbits manifest only a certain tendency towards it, while in cats and dogs it is fully absent.

In monkeys the formation of a new link in a chain of movements proceeds very rapidly—after only two combinations. It should be noted, however, that the stabilization of the whole chain takes place after 2–59 combinations, i.e., in a number of cases it lasts as long as in other mammals (Shirkova & Verevkina, 1960). In man, chains of motor reflexes are formed with striking speed. After the experiments the subjects said that they at once grasped the method of solving the task as stipulated in the preliminary instructions.

A conditioned inhibitor can be established to a chain of motor reflexes in man and animals. This means that if the conditioned stimuli of a basic chain of movements[10] are applied during the action of an indifferent agent and the chain of reflexes arising in response to them is not reinforced, then after a series of such negative combinations the animal or man ceases to react to the conditioned stimulation. In other words, a definite agent turns in a signal for the absence of reinforcement, i.e., becomes inhibitory.

Table 15 shows that phylogenetic distinctions cannot be shown in rabbits, cats, and dogs on the basis of the rate of formation of a conditioned inhibitor. Such an inhibitor is established much more rapidly in lower monkeys (macaques, baboons) and very rapidly in man.

It was further found that the exclusion of this inhibitory signal may serve as a basis for the formation of a new chain of motor reflexes. For example, a human subject or an animal performs an accidental or induced movement; at the same time the conditioned signals which must evoke definite movements leading to the accomplishment of the given aim begin to act (in the case of the animal it is food, while in the case of man it is the emergence of a signal confirming the correctness of the reaction); under these conditions the exclusion of the inhibitory agent turns into a factor reinforcing certain new movements. Thus it is possible to establish a single or chain motor reflex which eliminates the conditioned inhibitor or disinhibits the basic chain of alimentary conditioned reactions.

TABLE 15

RATE OF FORMATION OF A CONDITIONED INHIBITOR TO A CHAIN OF MOTOR REFLEXES

Animals	Minimum and maximum number of combinations
Pigeon	21–48
Rabbit	18–68
Cat	36–60
Dog	15–84
Monkey (macaque, baboon, chimpanzee)	2–12
Man	2–6

Such a disinhibitory chain of movements, consisting of two or three components, could be established by us in all animals except tortoises. As a result of our persistent attempts to establish similar reactions in tortoises, the latter ceased to react to the conditioned stimuli altogether.

Our investigations disclosed a definite process of formation of disinhibitory chains of movements in man and animals based on the exclusion of the conditioned inhibitor; this process is in principle similar to the process of formation of chains of

[10] By this term we designate a chain of movements which in animals is reinforced directly by food and in human subjects by beforehand stipulated signals.

TABLE 16
RATE OF FORMATION OF LINKS OF A DISINHIBITORY CHAIN OF MOVEMENTS
(Average figures)

Animals	Number of combinations required for the formation of:		
	The first link of a chain of movements	The second link of a chain of movements	The third link of a chain of movements
Pigeon	75	84	120
Rabbit	28	31	50
Dog	18	28	—
Monkey	14	19	26
Man	6	2	1

movements which are directly reinforced by food.

Table 16 shows that the rate of formation of a disinhibitory chain of movements steadily increases from pigeons to man.

As to phylogenetic distinctions in the rate of formation of separate links in a disinhibitory chain of movements (similar to that which was revealed during the formation of a chain of movements directly reinforced by food) they were not so strongly pronounced.

It may be said that in pigeons a three-component chain of movements was formed with great difficulty; with the "addition" of each new link, the stabilization of the chain became more and more impeded.

This lawful phenomenon was also manifested in mammals. In monkeys in which disinhibitory chains of movements are usually rapidly formed, considerable difficulties were sometimes observed. Only in man, who "guesses" or "grasps" the principle of solving a given task, the formation of each subsequent link of the chain of disinhibitory movements did not require any effort. A complex system of motor reflexes established in this manner has a common reinforcement (food in the case of animals, and a stipulated signal in the case of man); besides, separate links and parts of this system have their own reinforcement. Thus, the disinhibitory chain of movements is reinforced by the exclusion of the conditioned-inhibitory agent, while each link of the basic and disinhibitory chain is reinforced by putting into action the signal of the subsequent movement.

It was found that in experiments with human subjects this complex system could easily be extended owing to the formation of conditioned inhibitors of second and third orders and of disinhibitory chains of movements consecutively eliminating these inhibitors (Voronin & Napalkov, 1960).

Figure 1 presents a schematic diagram of the structure of such a system of motor conditioned reflexes in man consisting of a basic (A_1-B_2-C_3-D_4) chain of movements, conditioned inhibitors to it (J_1-J_2-J_3), and disinhibitory chains of movements (M_5-M_6-O_7, L_8-R_9-S_{10}, and V_{11}-W_{12}), as a result of which the conditioned inhibitors are eliminated. The experiment proceeds as follows: the subject is given the task of switching on an electric bulb (P); for this purpose, the subject, acting on his previous experience, performs a chain of movements (A_1-B_2-C_3-D_4); but if the conditioned inhibitors are in action (bulbs of other colors situated in definite points of the signal board), these movements do not lead to the solution of the task.

The subject switches on the conditioned stimuli in the same sequence in which they were established. He performs first the third chain of disinhibitory movements (V_{11}-W_{12}), the second (L_8-R_9-S_{10}) and, finally, the first (M_5-N_6-O_7). After that he performs the basic chain of movements (A_1-B_2-C_3-D_4), and thus finally solves the task, i.e., switches on the bulb (P).

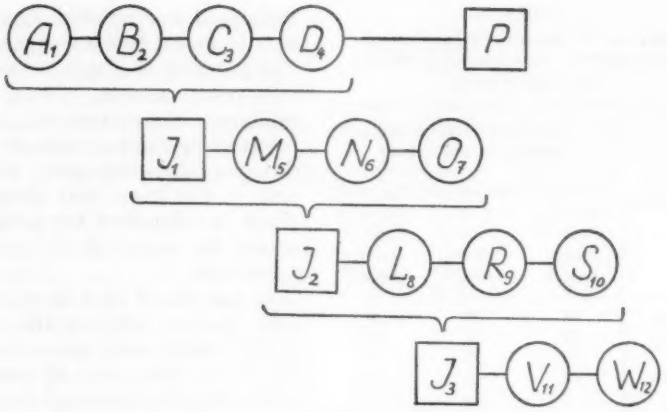


FIG. 1. Scheme illustrating the system of motor reflexes. (Explanations to be found in the text.)

We tested the possibility of establishing similar conditioned inhibitors and disinhibitory chains in representatives of our "series" of vertebrates. We found that a disinhibitory chain of movements is not formed in tortoises at all. In pigeons, rats, and dogs such a chain can be formed, but it is only of a primary character. In these animals it proved impossible to establish a conditioned inhibitor to a disinhibitory chain of movements and to form, on the basis of its exclusion, a second disinhibitory chain, i.e., a phenomenon of a secondary character.

The experiments of Shirkova and Verevkina (1960) showed that these temporary connections can be formed in monkeys (macaques, baboons, chimpanzees). This task, however, proved to be extremely difficult for monkeys, owing to which some exhibited neurotic states in the form of a disturbance of the entire conditioned-reflex activity. As stated above, conditioned inhibitors of third order are established in man; judging by their rapid formation, they do not constitute the limit of the subject's possibilities.

After the experiments we used to ask the subjects how they had managed to solve the given task; they answered that they had guessed that it was necessary consecutively to switch off those signals during the action of which the task could not be solved, i.e., the red light could not be switched on. It is quite clear that the subject was guided by the principle which underlies the structure of the system of signals and which was grasped by him in the course of experimentation; owing to the subject's capacity to generalize facts, he can solve even more complex tasks on the basis of this principle. In animals, owing to the absence of speech or of the "second signaling" system (according to the terminology of Pavlov), such generalization is impossible. The animal cannot therefore, "guess" the principle on which the structure of the system of signals, as well as of the reaction, is based; consequently, it cannot make use of this principle when solving the same but somewhat more complicated task.

It is noteworthy that monkeys differ greatly in this respect from other animals. They were able to solve more complex tasks, apparently, not because of their capacity to guess in

the way man does, but by virtue of their more developed "first signaling system," i.e., their capacity to impress more complex systems of temporary connections called forth by concrete (and not abstract) stimuli.

Owing to the highly developed capacity of monkeys to synthesize complex systems of external and internal stimuli, their concrete or object thinking differs greatly from that of other animals. At the same time this difference does not have a qualitative character; it is based on the capacity to form a wider range of temporary connections and on a higher degree of development of the integrative function of the brain. These capacities, which arose in our ancestors as far back as the prehistoric period, apparently, constituted the physiological precondition for the development of abstract thought.

Concerning the Mechanisms of Chain Motor Reactions

Two kinds of chain reactions studied at our laboratory are experimental models of those phenomena which are called "motor habits" in everyday life. Numerous physiological and psychological researches have been devoted to the study of these habits, and yet up to now no common view of their mechanisms has been reached. Some researchers refer to the reflex theory of Sechenov and Pavlov and believe that the function of voluntary movements is subject to the laws of conditioned reflex and that any elementary or complex habit is a partial case of this reflex. Other researchers positively deny the possibility of bringing these phenomena of acquired motor activity within the limits of the principle of temporary connections. Owing to this, they regard motor habits as phenomena qualitatively different from motor conditioned reflexes.

We adhere to the first point of view, and therefore are not inclined to oppose the psychological notion of motor habit to the physiological notion of motor conditioned reflex. Essentially these two notions are identical denoting one and the same phenomenon, but from different aspects. The physiological aspect concerns mainly the nervous mechanisms of this phenomenon, while the psychological aspect emphasizes its integrity and its general significance for the organism.

It is quite obvious that both above mentioned cases of chain motor reactions result from the consecutive combination of motor acts into a single chain of movements, i.e., from the analytico-synthetical process. It is likewise obvious that the formation of a motor reflex to an external stimulus in an animal is a phenomenon of analysis of exteroceptive and proprioceptive stimuli and of their synthesis into a single functional "combination center."

In man this functional combination is of a more complex character. It includes also the functional combination center which is formed in the area of the second signaling system. At the same time these "centers" are created in both signaling systems according to the same principles. Whereas in one case it is an external, directly acting agent; in the other—it is a word heard or seen, pronounced mentally or aloud, a word which designates the given agent.

In the first signaling system there takes place a combination of excitations coming from an external agent and from movements of the skeletal muscles, while in the second signaling system it is the excitations coming from speech and from the movements of the speech-motor apparatus that become fused.

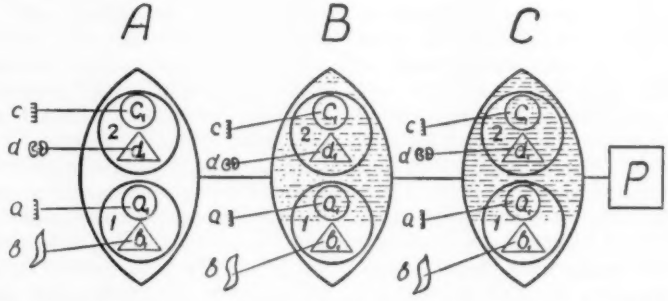
As a result of the synthesis of separate motor acts into a chain of movements (each of these acts arises in re-

sponse to a definite stimulus) and after their training this chain reaction is evoked by the action of the first stimulus alone. Such a motor set or automatized chain of movements is, as it were independent of all other chain stimuli; in man it is also independent of the second signaling system.

This phenomenon is, apparently, accounted for by the fact that as a result of constant training, excitation concentrates in corresponding nervous pathways according to the principle of negative induction and inhibits the "exteroceptive" and second signaling excitation (Figure 2). At the same time the action of the first stimulus and its reflection in the second signaling system acquire a "starting" faculty and these are not subjected to inhibition.

Inhibition, this arising in the structure of the motor reflex, disappears at once if there emerges an orienting reflex evoked, for example, by a violation of the fixed application of the stimuli. This explains why in the presence of unusual signals on the control panel a human subject stops his stereotyped movements in order to find out the cause of such a violation in the sequence of signals.

When a chain motor set is formed by consecutive addition of new links of a chain of stimuli and movements to links already established, the significance of the external stimuli is more pronounced as it is almost constantly observed. This is, apparently, explained by the fact that in this case the external stimuli play a double role: they reinforce the movement already effected, at the same time signaling the subsequent movement. As to stimulations of the second signaling system, they become more or less inhibited with the training of the chain of movements; when performing his movements, the subject is guided predominantly by exteroceptive and kinesthetic stimulations. These two kinds of stimulations may not be reflected in the second signaling system until their usual sequence is changed and an orienting reflex emerges (Ivanova, 1960). This influence of the orienting reflex testifies to the fact that the exteroceptive and kinesthetic stimuli were reflected in the second signaling system at the beginning of the formation of a chain



FIG, 2. Scheme illustrating an automatized chain motor reflex. (ABC combination "centers" of the links of a chain of conditioned reflexes. P = reinforcement of the chain of movements. 1. Combination centers of "first-signaling" system excitation: a = exteroceptors, a_1 = cortical focus of exteroceptive excitation; b = proprioceptors, b_1 = cortical focus of proprioceptive excitation; c = verbal stimulus, c_1 = cortical focus of verbal stimulation; d = proprioceptors of speech organs, d_1 = cortical focus of kinesthetic stimulations arising in speech organs. 2. Combination centers of "second-signaling" system excitation.)

of motor reflexes; but later with the training of the reflexes, they became inhibited.

The Concept of Reinforcement

In connection with the question of the mechanisms controlling the formation of chain motor reflexes it is important to have a clear idea of what we imply by the concept "reinforcement" and what its physiological mechanism is.

In our experiments with animals food served as a reinforcement, while in experiments with human subjects this role was played by a signal which confirmed the correct solution of the task. In both cases, in the course of formation of a chain of reflexes each preceding movement was reinforced by the stimuli of the subsequent movement. According to the point of view adhered to in Pavlov's laboratories, the coupling of a temporary connection results from the blazing of a nervous pathway in the place where the irradiating excitations from the conditioned and unconditioned cortical centers meet. This meeting of the irradiating excitations is possible if both centers are activated simultaneously. Then the conditioned excitation can reach the cortical representation of the unconditioned center only when a focus of excitation arises in it which according to this point of view, plays a reinforcing role.

It may be assumed that any kind of excitation—conditioned, unconditioned, orienting, alimentary, defensive, etc.—has a definite physicochemical expression. This assumption agrees with Pavlov's point of view advanced by him in his lectures as far back as 1912, and also in a number of his subsequent articles, according to which the quality or property of the stimulus determines the character of the nervous process or, in Pavlov's terminology, the "specific course of stimulation."

Although at present we do not possess any factual data concerning the physicochemical foundations of the innermost mechanisms involved in the coupling of temporary connections, it does not mean that we are unable to form certain ideas of these mechanisms by noting the results of experimental conditions on the organism and its responses to them. We may, for example, assume that the brain models the external influence in conformity with the various parameters of this influence, such as its strength, duration, frequency, etc. (Voronin & Sokolov, 1960; Sokolov, 1960). The nervous model is created by the group of neurons which store information concerning the properties of the stimulus. If, for example, the orienting reflex has been extinguished, but one of the parameters of its stimulus has changed, then the nervous model, formed as a result of the excitation of the reflex, does not coincide with the new supply of information. Owing to this noncoincidence, there arise impulses of "Discoordination" spreading over the descending cortico-reticular pathways evoking an orienting reflex which "washes off" the previously formed nervous model. Subsequent events develop according to the law of the orienting reflex; in case of its extinction (as a rule, this extinction is inevitable) a new nervous model is formed and it exists as long as the conditions of its emergence persist.

The mechanism of formation and reinforcement of a conditioned reflex may be presented in the following way: when two excitations of different codes coincide in time, one of them, which is the stronger reshapes the physicochemical structure of the other, and this leads to the creation of a single system of excitation. This system persists as long as the excitation coming along the afferent pathways can flow coordinatedly into the

stream of nervous impulses which are of a related nature. If this does not occur, the nervous process becomes inhibitory.

The process of automation of chain motor reactions takes place as a result of constant training of the nervous connections in the area of the kinesthetic analyzer; this ensures an unimpeded irradiation of the nervous process in them, irrespective of the nervous connections of other analyzers entering the structure of the given system of a complex reflex. For some reason this predominance of kinesthetic stimulation evokes a state of slight inhibition in the area of the second-signaling system connections, as well as connections between the movements and direct, first-signaling system stimulations; at the slightest violation of the system of stimuli this inhibition is washed off by the orienting reflex that arises. Hence, two states of the nervous processes—excitation and inhibition—are created in each of the "functional combination centers" which are presented in our scheme (Figure 2). It is the equilibrium of these two states which underlies the development of an automatized chain motor reflex.

General Conclusions

Our comparative-physiological investigations of higher nervous activity now permit us to draw the following two general conclusions, First, the peculiarities of higher nervous activity in vertebrates of different phylogenetical levels are based on the quantitative growth and complication of the conditioned reflex mechanisms; but the principle of the organism's interaction with the environment is identical in all cases. Individual experience and individual training of both the lower and the higher vertebrates are due to the establishment of connections between the concrete influences of the environment and the various activities of the organism. The number of these connections increases with the development of the nervous system, their structure assuming a more and more complex nature. Considering the same process from a psychological point of view, we may say that concrete or object thinking of animals becomes enriched with the development of their nervous system.

Even in the case of monkeys whose higher nervous activity is so greatly developed, there are no grounds for assuming the presence of any human-like mentality in them, or of any rudiments of such mentality. It may rather be said that human mental activity has preserved certain "remainders" of the animal intellect in the shape of concrete, image bearing, object thought but this is usually, dominated by abstract, verbal thought.

Second, the basic mechanism of higher nervous activity, which developed in the process of evolution of the animal world, is a mechanism of analysis and synthesis of stimulation falling on the organism, as well as of the activity arising in response to it.

On the basis of the factual material obtained by myself and by other researchers I am able to mark out three stages in the evolutionary development of higher nervous activity of animals (Voronin, 1960a). The distinctive features of each of these stages are: the state of perfection of the analytico-synthetical mechanism of higher nervous activity and correlation between its acquired and inborn components.

The first stage includes the higher nervous activity of animals possessing a primitive and coarsely differentiated nervous system. The behavior of these animals, just as that of present-day coelenterata, worms, and arthropoda, was determined predominantly by inborn reactions. At this

stage the role of reactions acquired during the individual life of the organism was insignificant. The general principle underlying the work of physiological mechanisms concerned with acquired activity does not, apparently, differ from the principle of temporary connections which are formed in modern animals. A low degree of development of the properties of the nervous processes and the capacity to retain nervous traces made possible formation only of elementary conditioned reflexes of a general nature such as locomotion in response to coarse signals which indicate circumstances favorable or unfavorable to the organism. This phase of development of higher nervous activity may be designated as a phase of simplest analysis and synthesis of elements of the organism's environment.

At the second stage of its development higher nervous activity was, apparently, of the same character as that in modern fish, amphibians, and reptiles. At this stage the analytico-synthetical mechanisms control more delicate individual relations between the organism and its environment, although the specific, hereditary mechanisms of these relations still play a predominant role. However, in this group of animals the range of temporary connections is greater, and this considerably increases the role of individual adaptive reactions. With the development of the locomotor apparatuses and their central coordination the local motor reactions occupy a more important place in the system of conditioned reflex activity. The development of the strength, equilibrium and, especially, lability of the excitatory and inhibitory processes enables the organism to modify its reactions within a comparatively short space of time in conformity with the changes that arise in the environment.

The third stage relates to that kind of higher nervous activity which is inherent in modern birds and mammals. A high degree of development of this activity in the ancestors of modern anthropoids and man played a particularly important role at this stage. These animals were distinguished from the rest of the animal world by their highly developed analytico-synthetical mechanisms, just as higher monkeys are distinguished by it at the present time.

At the third stage of evolutionary development of the higher nervous activity an ever-increasing role in the behavior of the organism is played by reactions acquired in the course of individual life.

At this stage not only hereditary-unconditioned reactions, but also reactions acquired in individual life cause the formation of temporary connections. Complexes of conditioned-unconditioned reflexes may originate complex multistage temporary connections. These nervous connections are utilized not only in the condtions in which they were formed, but also in new more or less similar conditions; in other words, the phenomenon of so-called "transfer of experience" takes place.

Thus, the third stage of development of higher nervous activity in animals is characterized by a high level of analytico-synthetical mechanisms, by the faculty of the neuronic apparatus to "impress" stimulations, as well as by a protracted retention of nervous traces, lability of the neurons, specialization of the analyzers and at the same time their integration. All this insured the finest individual adaptation of the organism (by means of temporary connections) to changing environmental conditions.

The above mentioned peculiarities of the third stage of development of higher nervous activity are more or

less manifest also at the second and even first stages. It may be said that the third stage includes the first and second ones, but differs from them, above all, quantitatively.

Comparing, for example, the results of our investigations on anthropoids and other vertebrates we can by no means state that there exists a specific qualitative difference between them. The point is that one and the same principle of individual adaptive reactions is of different significance on various levels of phylogenesis.

At the third stage a high degree of development was reached by that aspect of higher nervous activity which can be considered as concrete or image bearing thought.

At the level of man the qualitatively new signaling mechanisms, which were designated by Pavlov as the second signaling system of reality, gave rise to a new type of thoght, namely, abstract thought. This type is indissolubly bound up with concrete, image bearing thought, on the basis of which it arises and over which it dominates; it is of primary importance in man's relationships with the conditions of his existence.

REFERENCES

BARU, A. V. K metodike issledovania dvigatel'nykh pishchevykh uslovnykh refleksov u ptits. [Methods of investigation of motor alimentary conditioned reflexes in birds.] *Trud. Inst. Fiziol. Pavlova*, 1953, **2**, 449–453. (a)

BARU, A. V. Uslovnoye tormozhenie u ptits (utok i kur). [Conditioned inhibition in birds (ducks and hens).] *Trud. Inst. Fiziol. Pavlova*, 1953, **2**, 454–469. (b)

BARU, A. V., BOLOTINA, O. P., PRAZDNIKOV, N. V., & CHERNOMORDIKOV, V. V. K sravnitel'noi fiziologii uslovnovo refleksa na vremia. [Comparative physiology of the conditioned reflex to time.] In, *Proceedings sixteenth conference on problems of higher nervous activity: Theses and abstracts of communications.* Moscow-Leningrad: Izvestiya Akademiya Nauk SSSR, 1953. Pp. 35–36.

BEZNOSIKOV, B. O. K voprosu o fiziologicheskom mekhanizme trenirovki ugasatel'novo tormozhenia. [Contribution to the question of the physiological mechanism involved in the training of extinctive inhibition.] *Fiziol. Zh. SSSR*, 1954, **40**(6), 653–660.

BIRYUKOV, D. A. (Ed.) *Issledovaniye po evolutsii nervoni deyatel'nosti.* [Research into the evolution of the nervous activity.] Leningrad: Medgiz, 1959.

BIRYUKOV, D. A. *Ekologicheskaya fiziologia nervnoi deyatel'nosti.* [Ecological physiology of the nervous activity.] Moscow-Leningrad: Medgiz, 1960.

BOGOMOLOVA, E. M., SAAKYAN, S. A., & KOZOROVITSKY, L. B. Podrashtel'nye uslovnye refleksy u ryb. [Imitative conditioned reflexes in fish.] *Trud. Sovyeshch. Fiziol. Ryb*, 1956, 36–44.

BOLOTINA, O. P. Dvigatel'nye uslovnye refleksy na vremia u obezian. [Motor conditioned reflexes to time in monkeys.] *Trud. Inst. Fiziol. Pavlova*, 1952, **1**, 196–204. (a)

BOLOTINA, O. P. Dvigatel'nye uslovnye refleksy na vremia u sobak. [Motor conditioned reflexes to time in dogs.] *Trud. Inst. Fiziol. Pavlova*, 1952, **1**, 29–34. (b)

BOLOTINA, O. P. Vliiani broma i kofeina na uslovyne refleksy na vremia u sobak. [Effect of bromide and caffeine on conditioned reflexes to time in dogs.] *Trud. Inst. Fiziol. Pavlova*, 1953, **2**, 52–63.

BYKOW, K. M., & SLONIM, A. D. *Issledovanie slozhno-reflektornoi deyatel'nosti zhivotnykh i cheloveka.* [Investigation of complex reflex activity in animals and man.] Moscow-Leningrad: Izvestiya Akademiya Nauk SSSR, 1960.

CHERNOMORDIKOV, V. V. Uslovnye tormozhenie u cherepakh. [Conditioned inhibition in tortoises.] *Trud. Inst. Fiziol. Pavlova*, 1953, **2**, 479–489.

CHUMAK, V. I. K voprosu o mekhanizme uslovnovo refleksa na otnoshenie razdrazhitelei. [Contribution to the question of the mechanism of a conditioned reflex to the relationship of stimuli.] *Zh. vyssh. nervn. Deyatel.*, 1957, No. 1, 126–133. (a)

CHUMAK, V. I. Uslovnye refleksy na otnoshenie razdrazhitelei u zolotykh karasei, golubei, i krolikov. [Conditioned reflexes to the relationship of stimuli in goldfish, pigeons, and rabbits.] *Zh. vyssh. nervn. Deyatel.*, 1957, No. 1, 114–125. (b)

CHUMAK, V. I. Vliianie pereryva v opytakh na uslovnye refleksy u karpov, golubei, i krolikov. [Effect of intervals in experiments with conditioned reflexes in carp, pigeons, and rabbits.] *Nauchn. Dokl. vyssh. Shkoly*, 1958, No. 1, 93–95.

FIRSOV, L. A. Uslovnoye tormozhenie u

primatov. [Conditioned inhibition in primates.] *Trud. Inst. Fiziol. Pavlova*, 1954, 2, 433–447.

FIRSOV, L. A. Dvigatel'nye uslovnye refleksy na tsepi razdrazheitelei u sobak. [Motor conditioned reflexes to chains of stimuli in dogs.] *Zu. vyssh. nervn. Deyatel.*, 1953, No. 6, 842–851.

FIRSOV, L. A. Dvigatel'nye uslovnye refleksy na tsepi razdrazhitelei u detionyshei shimpanze. [Motor conditioned reflexes to chain stimuli in the young of chimpanzees.] *Zh. vyssh. nervn. Deyatel.*, 1955, No. 2, 247–254.

IORDANIS, K. A. Analiz slozhnykh uslovnoreflektornykh dvizhenii. [Analysis of complex conditioned-reflex movements.] *Nauchn. Dokl. vyssh. Shkoly*, 1958, No. 3, 78–83.

IORDANIS, K. A. Obrazovanie tsepei dvigatel'nikh uslovnykh refleksov iz chetyriokh zven'iev u obez'ian. [Formation of four-link chains of motor conditioned reflexes in monkeys.] *Vop. exp. Patol.*, 1959, 84–96. (a)

IORDANIS, K. A. Sravnitel'no-fiziologicheskie dannye ob uslovnom tormozhenii i uslovnom rastormazhivanii. [Comparative-physiological data on conditioned inhibition and conditioned disinhibition.] *Zh. vyssh. nervn. Deyatel.*, 1959, No. 1, 126–134. (b)

IVANOVA, V. I. Formirovanie tsepnykh dvigatel'nykh uslovnykh refleksov u ryb, golubei, i krolikov. [Formation of chain motor conditioned reflexes in fish, pigeons, and rabbits.] *Nauchn. Dokl. vyssh. Shkoly*, 1960, No. 3, 86–90.

KOGAN, A. B. *Osnova fiziologii vysshei nervnoi dyatel'nosti.* [Fundamentals of the physiology of higher nervous activity.] Moscow: Vysshey Shkoly, 1959.

KOZOROVITSKI, L. B. Podrazhatel'nye uslovnye refleksy u shimpanze. [Imitative conditioned reflexes in chimpanzees.] In *Proceedings seventeenth conference on problems of higher nervous activity: Theses of communications.* Moscow-Leningrad: Izvestiya Akademiya Nauk SSSR, 1956. Pp. 64–65.

KRUSHINSKAYA, N. L., KHOLODOV, Y. A. SHURANOVA, ZH. P., & SHCHERBINA, Z. D. Sravnitel'no-fiziologicheskie dannye o differentsirovanii dvukh polozhitel'nykh razdrazhitelei. [Comparative-physiological data on the differentiation of two active stimuli.] In, *Proceedings nineteenth conference on problems of higher nervous activity: Theses and abstracts of communications.* Moscow-Leningrad: Izvestiya Akademiya Nauk SSSR, 1960. Pp. 188–189.

MAKOKINA, S. M., & KHOLODOV, Y. A. Uslovnoye tormozhenie i uslovnoye rastormazhivanie u shimpanze, pavianov i sobak. [Conditioned inhibition and conditioned disinhibition in chimpanzees, baboons, and dogs.] *Zh. vyssh. nervn. Deyatel.*, 1959, No. 1, 555–560.

MALINOVSKY, O. V. Metodika dvigatel'nykh pishchevykh uslovnykh refleksov u krolikov. [Methods of motor alimentary conditioned reflexes in rabbits.] *Fiziol. Zh. SSSR*, 1952, No. 5, 637–639. (a)

MALINOVSKY, O. V. Uslovnye refleksy vtorovo poryadka u obez'ian. [Conditioned reflexes of second order in monkeys.] *Trud. Inst. Fiziol. Pavlova*, 1952, 1, 205–212. (b)

MALINOVSKY, O. V. Uslovnoye tormozhenie u krolikov. [Conditioned inhibition in rabbits.] *Trud. Inst. Fiziol. Pavlova*, 1953, 2, 470–478. (a)

MALINOVSKY, O. V. Vyrabotka vremennoi svyazi na indifferentnye razdrazhiteli u krolikov. [Elaboration of a temporary connection to indifferent stimuli in rabbits.] *Trud. Inst. Fiziol. Pavolva*, 1953, 2, 335–340. (b)

NAPALKOV, A. V. Fiziologicheskie mekhanizmy lezhaschchie v osnove formirovania tsepei dvigatel'nykh uslovnykh refleksov. [Physiological mechanisms underlying the formation of chains of motor conditioned reflexes.] *Nauchn. Dokl. vyssh. Shkoly*, 1958, No. 2, 66–73. (a)

NAPALKOV, A. V. Izuchenie zakonomernostei vyrabotki slozhnykh sistem uslovnykh refleksov. [Study of the laws governing the elaboration of complex systems of conditioned reflexes.] *Vestn. Moskovsk. U.*, 1958, No. 2, 75–83. (b)

NAPALKOV, A. V., & VEREVKINA, G. L. Irradiatsia vozbuzhdenia iz pit'ievovo tsentra po raneye vyrabotannoi pishchevoi tsepi refleksov. [Irradiation of excitation from the thirst center over a previously elaborated alimentary chain of reflexes.] *Vop. exp. Patol.*, 1959, 52–61.

OVCHINNIKOVA, N. P. Dvigatel'nye pishchevye uslovnye refleksy na tsepi razdrazhitelei u ptits. [Motor alimentary conditioned reflexes to chains of stimuli in birds.] *Zh. vyssh. nervn. Deyatel.*, 1955, No. 1, 87–96.

PAVLOV, I. P. *Lektsii o rabote bol'shikh polusharii golovnovo mozga.* [Lectures on the work of the cerebral hemispheres.] Moscow: Gosizdat, 1926.

PAVLOV, I. P. *Dvattsatiletnii opyt ob'ektivnovo izuchenia vysshei nervnoi deyatel'nosti (povedenia) zhivotnykh.* [Twenty years experience in the objective study of higher nervous activity (behavior) of animals.] Moscow: Biomedgiz, 1938.

POLIVANNAYA, M. F. K voprosu o trenirovke svoistv nervnykh protsessov. [Contribution to the question concerning the training of the properties of nervous processes.] *Zh.*

vyssh. nervn. Deyatel., 1960, No. 4, 620–625.

PRAZDNIKOVA, N. V. Metodika issledovania dvigatel'nikh pishchevykh uslovnykh refleksov u ryb. [Methods of investigation of motor alimentary reflexes in fish.] *Zh. vyssh. nervn. Deyatel.*, 1953, No. 3, 464–468. (a)

PRAZNIKOVA, N. V. Pishchevye dvigatel'nye uslovnye refleksy i uslovny tormoz u ryb. [Alimentary motor conditioned reflexes and conditioned inhibition in fish.] *Trud. Inst. Fiziol. Pavlova*, 1953, 2, 370–383. (b)

PRAZDNIKOVA, N. V. Pishchevye dvigatel'nye uslovnye refleksy na tsepi razdrazhitelei u ryb. [Alimentary motor conditioned reflexes to chains of stimuli in fish.] *Zh. vyssh. nervn. Deytael.*, 1955, No. 6, 901–911.

PRAZDNIKOVA, N. V., & FIRSOV, L. A. Vzaimootnoshenie dvigatel'nova i golosovovo komponentov uslovnoi reaktsii u obezian (kaputsinov) pri peredelke odnovo iz chlenov tsepnovo razdrazhitelya. [Interrelation of the motor and vocal components of a conditioned reaction in monkeys (capuchins) during the reversal of one of the components in a chain stimulus.] *Trud. Inst. Fiziol. Pavlova*, 1953, 2, 306–315.

ROKOTOVA, N. A. Obrazovanie vremennykh svyazei u sobak pri deistvii sledov indifferentnykh razdrazhenii. [Formation of temporary connections in dogs under the action of traces of indifferent stimuli.] *Trud. Inst. Fiziol. Pavlova*, 1952, 1, 35–42.

ROKOTOVA, N. A. O vremennykh svyasyakh na indifferentnye razdrazhiteli u antropoidon (shimpanze). [Temporary connections to indifferent stimuli in anthropoids (chimpanzees).] *Trud. Inst. Fiziol. Pavlova*, 1953, 2, 289–294.

ROKOTOVA, N. A. K Voprosu o fiziologicheskikh mekhanizmakh Vremennykh svyazei na indifferentnye sazdrazhiteli. [Contribution to the question of the physiological mechanisms of connections to indifferent stimuli.] *Zh. vyssh. nervn. Deyatel.*, 1954, No. 4, 516–525. (a)

ROKOTOVA, N. A. O metodike opredelenia tipa nervnoi sistemy u cheloveka. [Methods of determining the type of nervous system in man.] *Fiziol. Zh. SSSR*, 1954, No. 6, 727–729. (b)

ROKOTOVA, N. A. Tsepnye dvigatel'nye uslovnye refleksy u sobak. [Chain motor conditioned reflexes in dogs.] *Zh. vyssh. nervn. Deyatel.*, 1954, No. 6, 833–841. (c)

ROKOTOVA, N. A. O fiziologicheskom mekhanisme razlichenia polozhitel'nykh uslovnykh signalov svyazannych s raznymi uslovnymi reaktsiami. [Physiological mechanism involved in the differentiation of active conditioned signals connected with various conditioned reactions.] *Zh. vyssh. nervn. Deyatel.*, 1955, No. 3, 385–392.

SECHENOV, I. M. Refleksy golovnova mozga. [Reflexes of the brain.] *Fiziol. nervn. Sis.*, 1952, 1.

SEVERTSEV, A. N. *Glavnoye napravlenie evolutsionnovo protsessa.* [The main direction of the evolutionary process.] Moscow: Biomedgiz, 1934.

SHIRKOVA, G. I. Dvigatel'nye uslovnye refleksy na tsepi razdrazhitelei u nizshikh obez'ian. [Motor conditioned reflexes to chain of stimuli in the lower apes.] *Biull. eksp. Biol. Med.*, 1949, No. 9, 174–178.

SHIRKOVA, G. I., & VEREVKINA, G. L. Dvigatel'nye tsepnye uslovnye refleksy u obez'ian. [Motor chain conditioned reflexes in monkeys.] *Dokl. Akad. Nauk SSSR*, 1960, No. 3, 730–734.

SOKOLOV, E. N. Nervnaya model stimula i orientirovochny refleks. [The nervous model of a stimulus and the orienting reflex.] *Vop. Psikhol.*, 1960, No. 4, 61–72.

TAGIEV, SH. K. Sintez dvukh razlichnykh dvigatel'nykh uslovnykh refleksov tsep dvizhenii u ryb, gloubei, i krolikov. [Synthesis of two different motor conditioned reflexes into a chain of movements in fish, pigeons, and rabbits.] *Zh. vyssh. nervn. Deyatel.*, 1958, No. 3, 431–436.

VORONIN, L. G. Dvigatel'nye uslovnye refleksy vyrabotannye tak nazyvayemym metodom mekhanicheskoi dressirovki. [Motor conditioned reflexes elaborated by the so-called method of mechanical training.] *Trud. Inst. Evol. Fiziol. Patol. Vyssh. Nervn. Deyatel.*, 1947, 1, 111–132. (a)

VORONIN, L. G. K voprosu obimitationnykh sposobnostyakh u nizshikh obez'ian. [Contribution to the question of the imitative capacities of the lower apes.] *Fiziol. Zh. SSSR*, 1947, 33(1), 373–380. (b)

VORONIN, L. G. Nekotorye dannye ob uslovnom reflekse na vremia u nizshikh obez'ian (pavianov, makakov, i martyshek). [Some findings on the conditioned reflex to time in the lower apes (baboons, macaques, long-tailed monkeys).] *Zh. vyssh. nervn. Deyatel.*, 1951, No. 1, 213–222.

VORONIN, L. G. *Analizis i sintezis slozhnykh razdrazhitelei u vysshikh zhivotnykh.* [Analysis and synthesis of complex stimuli in higher animals.] Leningrad: Medgiz, 1952.

VORONIN, L. G. Ob evolutsii svoistv nervynkh protsessov. [Evolution of the properties of the nervous processes.] Communications delivered at the nineteenth International Congress of Physiologists, Moscow, 1953.

VORONIN, L. G. Nekotorye itogi sravnitel'no-fiziologicheskovo izuchenia vysshei nervnoi deyatel'nosti. [Some results of the comparative-physiological study of the higher ac-

tivity.] *Izv. Akad. Nauk SSSR, Ser. Biol.*, 1954, No. 5, 122–134. (a)

VORONIN, L. G. Yeshcho raz o skorosti obrazovania uslovnykh refleksov. [Once more about the rate of formation of conditioned reflexes.] *Zh. vyssh. nervn. Deyatel.*, 1954, No. 5, 756–767. (b)

VORONIN, L. G. O nekotorhkh voprosakh sravnitel'noi fiziologii vysshei nervnoi deyatel'nosti. [Some questions of the comparative physiology of the higher nervous activity.] *Vestn. Moskovsk. U.*, 1955, No. 4–5, 207–217.

VORONIN, L. G. *Sravnitel'naya fiziologia vysshei nervnoi deyatel'nosti.* [Comparative physiology of the higher nervous activity.] Moscow: Izvestiya Moscow University, 1957.

VORONIN, L. G. Sravnitel'no-fiziologicheskie dannye k probleme vozniknovenia vysshei nervnoi deyatel'nosti cheloveka. [Comparative-physiological data relating to the problem of the origin of the higher nervous activty.] *Vop. Antropol.*, 1960, No. 4, 21–23. (a)

VORONIN, L. G. *Sravnitel'no-fiziologicheskie dannye o tsepnykh dvigatel'nykh uslovnykh refleksakh: Evolutsia fiziologicheskich funktsii.* [Comparative-physiological data on chain motor conditioned reflexes: Evolution of physiological functions.] Moscow-Leningrad: Izvestiya Akademiya Nauk SSSR, 1960. (b)

VORONIN, L. G., & IORDANIS, K. A. K sravnitel'no-fiziologicheskomu analizu slozhnykh uslovnoreflektornykh dvizhenii. [Comparative-physiological analysis of complex conditioned reflex movements.] *Nauchn. Dokl. vyssh. Shkoly*, 1960, No. 1, 59–67.

VORONIN, L. G., & IORDANIS, K. A. O vzaimootnoshenii tormoznovo i razdrazhitel'novo protsessov pri slozhnoi uslovnoreflektornoi deyatel'nosti. [Interrelation of the inhibitary and excitatory processes in complex conditioned reflex activity.] *Zh. vyssh. nervn. Deyatel.*, 1961, No. 1, 99–105.

VORONIN, L. G., & NAPALKOV, A. V. Metodicheskie priomy obrazovania slozhnykh sistem dvigatel'nykh uslovnykh refleksov u zhivotnykh. [Methods of formation of complex systems of motor conditioned reflexes in animals.] *Zh. vyssh. nervn. Deyatel.*, 1959, No. 5, 788–791.

VORONIN, L. G., & NAPALKOV, A. V. K metodike izuchenia vysshei nervnoi deyatel'nosti cheloveka. [Methods of studying the higher nervous activity of man.] *Dokl. Akad. Ped. Nauk RSFSR*, 1960, 2, 95–89.

VORONIN, L. G., & SOKOLOV, E. N. Cortical mechanisms of the orienting reflex and its relation to the conditioned reflex. Paper presented at Moscow Colloquium on Electroencephalography of Higher Nervous Activity, Moscow, 1960.

VORONIN, L. G., & SHIRKOVA, G. I. O vliyanii dlitel'novo pereryva v opytakh na dvigatel'nye uslovnye refleksy u obez'ian makaki-resusa. [Effect of a long interval in experiments with motor conditioned reflexes in monkeys (M. Rhesus).] *Trud. Sukhumsk. med.-biol. Stants. AMN SSSR*, 1949, 1, 67–79.

(Received May 18, 1961)

Psychological Bulletin
1962, Vol. 59, No. 3, 196–210

THE PERSONALITY OF EPILEPTICS: A DISCUSSION OF THE EVIDENCE[1]

BARBARA TIZARD
Institute of Psychiatry, Maudsley Hospital, London

The aim of this article is to assess the extent to which the various theories about the personality of epileptics have been affirmed or refuted, and to discuss some of the more important variables which should in the future be investigated by those working in this field. The complexity of the theoretical and methodological problems involved is also pointed out.

THEORIES

The theories discussed here are for the most part generalizations of particular clinicians' experiences, with little attempt to relate them to a wider body of knowledge. Five basic theories may be distinguished as follows:

1. That all or most epileptics share a charactristic personality. This was the original theory advanced in the nineteenth century by Falret and Féré, and extensively held on the Continent. It still has advocates, especially in Germany. The personality traits considered characteristic of epileptics have been variously described by different authors. There is some agreement on a basic syndrome of perseveration and viscosity in both the intellectual and affective spheres. Emotional explosiveness has also been named as a central trait by most authors, either co-existing with viscosity (Minkowska, 1946) or replacing it (Clark, 1918; Stauder, 1938). Traits such as suspiciousness, religiosity, meticulousness, and selfishness have been emphasized by others. According to these writers, the epileptic personality and the predisposition to convulsions are constitutionally determined. American writers who have recognized a characteristic if not universal epileptic personality have usually attributed it to the frustrating environment and the social stigma to which the epileptic is exposed, or to the effect of brain dysfunction on the personality, or to both these factors (Notkin, 1928; Revitch, 1955).

2. That there is no characteristic epileptic personality, and the same range and combination of personality traits may be found among epileptics and nonepileptics (Lennox, 1944). This theory was widely held in America during the "thirties" and "forties," particularly by clinicians whose main experience was with private patients.

3. That there is no characteristic epileptic personality or personality disturbance, but a higher proportion of neurotic disturbance is found among epileptics than among nonepileptics (Bridge, 1949).

4. That there is no characteristic epileptic personality, but epileptics tend to have a personality resembling that of patients with organic lesions, which differ from that of normal persons. This theory has been particularly favored by clinicians working with epileptic children (Bradley, 1951).

Each of these theories has been advanced largely on the basis of clinical observation, but in each case has been supported by evidence from

[1] In receipt of a grant from the Medical Research Council.

I am indebted to D. A. Pond and J. H. Margerison for many helpful discussions.

more objective studies, which will be considered in the sections on Rorschach Studies and Studies Other than Rorschach Studies.

5. That there is no characteristic personality common to all or most epileptics, but different types of personality are associated with different types of epilepsy. This is not a new theory—as far back as 1938 German authors were describing two distinct personality types, associated with idiopathic and symptomatic epilepsy. Since about 1950, however, with the delineation of, and growing interest in temporal lobe epilepsy, it has assumed a new importance. Temporal lobe epilepsy is now generally held to be associated with personality disturbance, and many of the unpleasant traits formerly attributed to epileptics in general are now said to characterize temporal lobe epileptics. In some neurosurgical centers temproal lobectomy is performed primarly in order to alleviate the personality disturbance of these epileptics. Some writers also describe a characteristic personality pattern found in children with petit mal epilepsy, and Gastaut and his colleagues have recently delineated an idiopathic epileptic personality type.

The theory that differences in personality among epileptics are related to differences in type of epilepsy has, in the case of temporal lobe epilepsy in turn been related to the results of experimental work with animals and to the results of surgery. Its advocates point out that bilateral ablation of the rhinencephalon in monkeys results in increased activity and docility (Kluver & Bucy, 1939) hence, the hypoactivity and aggressive outbursts of the patient with temporal lobe epilepsy may be considered to represent a state of excitation of the rhinencephalon. Moreover, some neurosurgical centers report a specific decrease in aggressive behavior after temporal lobectomy for epilepsy (Alajouanine, Nehlil, & Houdart, 1958; James, 1960). This theory is thus more than a clinical generalization, and has a rather different status from those outlined above. For this reason, as well as for its therapeutic implications, and because it has received little attention in psychological journals, the clinical evidence on which it is based will be briefly described. The evidence from psychological tests will be discussed in the sections on Rorschach Studies and Studies Other than Rorschach Studies.

The clinical evidence includes incidental observations, systematic studies of series of patients, and observations of the large proportion of patients with temporal lobe epilepsy amongst epileptics in mental hospitals. Typical of the incidental observations is Robertiello's (1953) description of the child with psychomotor epilepsy as "usually cooperative, good, quiet and overcontrolled, but with episodes of impulsive and often violent and destructive behavior." Peterman (1953) describes him as a child with "an abnormal personality, with recurring episodes of behaviour disorder," in contrast to the child with petit mal epilepsy, who is usually "mentally precocious, alert, sensitive and temperamental." Pond's contrast is between the aggressive child with temporal lobe epilepsy, and the timid passive child with petit mal epilepsy (Pond, 1952, 1961).

More systematic studies have reported the incidence of personality disturbance among temporal lobe epileptics. Hill (1957) reported that 50% suffer from personality disorders, and of these 25% have psychotic episodes. Gibbs (1958) stated that about 40% of the patients with psychomotor epilepsy have severe personality disorders, and of these about

one third are classifiable as psychotic. Gastaut and Gastaut (1951 unpublished) found that 52% of psychomotor epileptics attending outpatient clinics have psychiatric disorders and Bingley (1958) identified personality changes in 52% of epileptics with temporal lobe foci. Vislie and Henriksen (1958), reviewing the psychiatric symptoms of 162 epileptics, excluding psychotics, found a tendency for a higher incidence of neurotic symptoms when the localization signs pointed to the temporal region. The only comparable study with children, separating personality disturbance from intellectual defect, is that of Glaser and Dixon (1956). They found that 19 out of 25 children with psychomotor seizures had interictal personality disturbance.

Two recent studies have reported a large proportion of patients with temporal lobe epilepsy among epileptics in mental hospitals. Liddell (1953) reported an incidence of 50% and Roger and Dongier (1950) 64%, as compared with 30% in a series of outpatient epileptics (Gastaut, 1950).

The nature of the personality disturbance has been variously described. Gibbs (1958) found no specific symptoms, but Mulder and Daly (1952) stressed the frequency of anxiety and depression. Falconer, Hill, Meyer, and Wilson (1958) found explosive or persistent aggressiveness the commonest symptoms, as did Liddell (1953) and Roger and Dongier (1950) in their surveys of epileptics with temporal foci in mental hospitals. Paillas (1958) particularly noted the frequency of slowness, adhesiveness, and perseveration as well as aggressiveness, and Bingley (1958) found the commonest syndrome was "adhesiveness in the intellectual, emotional and volitional spheres." Vislie and Henriksen (1958) however found these traits were associated with evidence of diffuse lesions and organic dementia rather than temporal localization.

There are two main difficulties in evaluating these studies. In the first place, there appear to be important differences between the epileptic populations studied. In some studies, mainly patients with pronounced psychiatric symptoms were included (Falconer et al., 1958; Liddell, 1953; Mulder & Daly, 1952; Roger & Dongier, 1950). In some studies the criterion for selection was purely clinical (Glaser & Dixon, 1956); in others, electrophysiological (Bingley, 1958; Vislie & Henriksen, 1958); whilst in still others a combination of clinical and electrophysiological criteria was used (Mulder & Daly, 1952; Gastaut et al., 1955). Since electroclinical correlations are far from perfect, different populations may have been sampled. This point is discussed further in Section 4.

Secondly, no study shows an adequate appreciation of the problems of bias and reliability in the judgments made. Precautions are not reported to prevent contamination of EEG interpretation by clinical data and the reliability of the EEG and clinical diagnoses is not assessed. Bingley's is the only study which describes reassessment of the EEG records by another judge, without knowledge of the previous interpretation of the patient's clinical symptoms. The position is worse in respect of the assessment of the presence or type of psychiatric disturbance. This is left undefined in all studies and is never made without knowledge of clinical status. Nor is the reliability of the judgments assessed.

One study of greater methodological sophistication than the others is that of Nuffield (1961). The EEG records of 233 epileptic children who had attended the Maudsley Hospital dur-

ing the previous 10 years were classified into 7 electrophysiological groups. Each child was rated for aggressiveness and neurotic manifestations on the basis of the answers which had been recorded on the standardized psychiatric case history to such questions as "Is he irritable, a bully, timid, sensitive?" The mean "aggressive" score derived from these ratings of the children with a temporal lobe EEG focus was higher than that of any other EEG group, while their mean "neurotic" score was the lowest. Conversely, the mean aggressive score of the 3/sec SW group was the lowest, and their mean neurotic score was the highest. The correlations between fit patterns and behavioral ratings were much lower than those between EEG classifications and behavioral ratings. In this study contamination between EEG data and behavioral ratings was avoided, but the reliability of the judgments was not assessed.

Rorschach Studies

The great majority of psychologists who have studied the personality of epileptics have used the Rorschach test, and the evidence from these studies will be considered first. A specific epileptic Rorschach protocol common to all or many epileptics has been described by some authors (Minkowska, 1946; Rorschach, 1942; Stauder, 1938). More cautiously, others have concluded that while there is no specific epileptic personality, epileptics share many traits and can be identified by the use of the Rorschach test (Altable, 1947; Bovet, 1936; Kelly & Marquilies, 1940; Piotrowski, 1947; Zehrer, 1951). Other authors, however, have found no statistically significant difference between the Rorschach scores and patterns of epileptic and nonepileptic groups (Kogan, 1947; Lisansky, 1948; Shaw & Cruikshank, 1957). The theory that the personality of epileptics resembles that of brain injured patients is supported by a number of Rorschach studies (Piotrowski, 1947; Ross, 1941; Zimmerman, Burgemeister, & Putnam, 1951). Other studies, however, report evidence of neurosis but no signs characteristic of brain injury (Arluck, 1941; Kaye, 1951; Richards, 1952; Zehrer, 1951). The belief that temporal lobe epileptics share a characteristic personality is supported by three Rorschach studies (Delay, Pichot, Lemperiere, & Perse, 1955; Gastaut, Morin, & Lesevre, 1955; Paillas & Subirana, 1950). The Rorschach pattern identified as characteristic is, however, quite different in each of these studies.

There is thus evidence from Rorschach studies for and against each theory about the personality of epileptics which has been advanced. In attempting to account for the contradictory nature of these findings, one is struck by two rather crude methodological errors that invalidate the great majority of the studies.

1. Many of the earlier studies used institutionalized epileptics, a very small and uncharacteristic sample of the epileptic population. These patients are mentally disturbed and may also be of low intelligence. All the descriptions of a specific epileptic Rorschach protocol derive from studies of such patients.

2. The great majority of studies have not controlled for IQ. They have compared the Rorschach protocols of epileptics with the Rorschach norms, assuming either that intellectual level was not an important factor, or that the epileptic group and the Rorschach normal group were of similar intelligence. However, evidence has recently accumulated which shows that intellectual level affects a

great range of Rorschach responses in all-pervasive ways. Neff and Lidz (1951) in a study of army personnel divided into three subgroups according to intelligence showed that only the group with superior intelligence gave Rorschach responses usually considered typical of normal persons. Men with average and below average intelligence gave responses that by the usual Rorschach norms would be considered indicative of emotional disturbance. Similar findings are reported by Wedemeyer (1954). Later, Neff and Glaser (1954) showed that in Beck's normal group, used to furnish control data for many Rorschach studies, the percentage of those with high school education was two and a half times greater than in the general population. The epileptic protocols have thus been assessed and found in various ways abnormal by comparison with those of a group of above average intelligence.

The same kind of evidence invalidates most of the Rorschach studies of children. Control groups have rarely been used, and the epileptic protocols have been compared with published Rorschach child norms, such as the Ames norms. Fielder and Stone (1956) have shown, however, that three quarters of the children in the Ames Rorschach normal group came from professional and managerial classes. The responses from their own group of predominantly lower class normal children departed very considerably from the Ames' criteria of normality.

It thus appears that the conclusions of the great majority of Rorschach studies are unacceptable. In fact, the Rorschach scores of the epileptic groups resemble, and compare rather favorably with, those reported by Neff and Lidz and Wedemeyer for normal service groups of equivalent IQ. Even if, however, one considers only those studies which compare epileptics living in the community with nonepileptics, matched for age and IQ, the results appear no less conflicting than before.

There are only five such studies published. In addition, two studies have compared the Rorschach protocols of temporal lobe epileptics with other epileptics. Lisansky (1948) compared the Rorschach protocols of 10 adult noninstitutionalized epileptics and 10 diabetics, matched for age, education, and duration of illness. Both groups were of average intelligence. She found that the only significant difference between them was the slower response time of the epileptic group. The Rorschach "epileptic signs" did not differentiate the two groups.

Piotrowski (1947) compared the records of 25 epileptic adults, "none psychotic, hospitalized, or conspicuously deteriorated" and 25 hysterics of similar age and IQ. He found 14 "epileptic signs." Six of these had previously been included in his scale for differentiating patients with organic lesions. Most of the other signs can be found in various scales for differentiating neurotic or other disturbed patients. He also found that the presence of seven or more of these signs in a protocol identified 80% of his epileptic group, while none of the hysterics' records had more than four signs.

Arluck (1941) compared the Rorschach records of 16 idiopathic epileptics without an EEG focus, aged 10–21, with control groups of their sibs, of cardiac patients, and of the sibs of cardiac patients. The groups were equated for sex, age, IQ, and socioeconomic status. He found several statistically significant differences

in the Rorschach scores. The epileptic group showed more signs of color shock, had fewer Ws in their records, and their mean time for response was greater. All three of these signs are included among Piotrowski's epileptic signs, and the first two occur in most Rorschach scales that claim to differentiate normals from any other groups.

Shaw and Cruikshank (1951), however, obtained negative results when they compared the Rorschach protocols of 25 institutionalized idiopathic epileptic children and 25 institutionalized nonepileptic children matched for age, sex, and IQ. The only significant differences found were in FV and number of different types of content, both being larger for the epileptic group.

Kogan (1947) compared the Rorschach protocols of 10 idiopathic epileptic children with behavior problems and 10 nonepileptic children attending the same child guidance clinic, matched for age, IQ, and severity of emotional disturbance. She found no statistically significant differences between the two groups on any Rorschach variable. Clinically, both groups contained shy, quiet, anxious children and aggressive, antisocial children.

Delay et al. (1955) studied the Rorschach protocols of 50 epileptics of average intelligence classified according to presumed etiology and also site of EEG focus, if any. They found that 48% of the whole group gave seven or more of Piotrowski's epileptic signs, and 68% gave five or more of his organic signs. Patients with "psychomotor temporal epilepsy" had significantly more extratensive Rorschach protocols than the rest. Gastaut et al. (1955) compared the Rorschach protocols of 60 noninstitutionalized psychomotor epileptics of average intelligence and a group of idiopathic epileptics. Only general impressions of the results were given. They reported that 72% of the psychomotor epileptics showed a syndrome of hypoactivity and emotional indifference or depression. It was the idiopathic group whose protocols were predominantly extratensive.

There is thus on the one hand Piotrowski's evidence that the Rorschach protocols of epileptics resemble each other but differ from those of hysterics by the presence of certain characteristic signs. This is partially supported by Arluck's study, and by Delay's finding that the Rorschach protocols of 48% of an epileptic group of average intelligence contain a large number of these signs. On the other hand three further studies have been unable to differentiate the Rorschach protocols of epileptics and normal children, disturbed epileptic children and disturbed nonepileptic children, and epileptic adults and diabetic adults. Two studies of the Rorschach protocols of patients of average intelligence with temporal lobe epilepsy reached opposite conclusions.

The extent to which the clinical status of the patients in these studies is comparable will be discussed below. The question must be raised, however, whether there are any inadequacies in the Rorschach test which make it an unsuitable tool with which to investigate the personality of epileptics. The primary concern of Piotrowski (1947) and many workers in this field was differential diagnosis. Yet the Rorschach test has repeatedly been shown to be an unsatisfactory diagnostic instrument. When the items originally found to differentiate two groups are put into the form of a scale and a cross-validation study is made by a different worker on a fresh

sample, the differentiating power of the scale invariably drops considerably (Yates, 1954). Subjective scoring systems and test unreliability no doubt contribute to this finding. The consistency of the subject, subject reliability, scoring reliability, interpretation reliability, and enquiry reliability have proved disappointingly low (Baughman, 1951; Campbell & Fiddleman, 1959; Fiske, 1959). Moreover, where real differences in personality are known to exist between groups the Rorschach test cannot be relied upon to detect them. Studies in the last decade which have taken care to avoid contamination of Rorschach interpretations with clinical data have found slight or no significant differences between the Rorschach scores of hospitalized schizophrenics and normals (Friedman, 1952), between different psychiatric groups (Wittenborn & Holzberg, 1951), or between neurotics and schizophrenics (Reiman, 1953). Too much confidence cannot therefore be placed in the Rorschach test as an instrument for determining whether epileptics as a group resemble or differ from neurotics, normals, or brain injured groups.

Some studies have used the Rorschach test as a tool not only to differentiate between epileptic and other groups, but to describe the characteristics of the epileptic personality. Thus Arluck (1941) deduced from the protocols of his group that they suffered emotional strain, had much conflict within their basic personality configuration, and tended to adjust by withdrawing from the external world.

Such a use of the test depends on the assumption that different modes of response to the blots are determined, in a manner that is known, by dominant and enduring personality traits. Thus an extratensive Rorschach experience balance (i.e., a record with overemphasis on color) is said to be characteristic of an egocentric emotionally explosive personality. However, as Pruyser points out (1953), Rorschach, in common with most of his contemporaries, believed that epileptics are by constitution predisposed to these traits, When he found an unusually large number of color responses, especially CF and C, in the protocols of deteriorated epileptics, he concluded that such responses were determined by emotional explosiveness. Alternative explanations were not considered or investigated. Recent experimental work, however, strongly suggests that there is no relationship between number and type of color responses and emotional lability (Baughman, 1954; Keehn, 1954; Lazarus & Oldfield, 1955).

Attempts to validate the alleged relationship between Rorschach signs and such traits as constriction and impulsiveness have been no more successful (Carp, 1950; Holtzman, 1950). Very few such studies have been made. There is no adequate evidence for most of the Rorschach sign-trait correlations, and interpretations such as Arluck's have to be accepted on faith. Hence, even if consistent findings had emerged from the Rorschach studies, their significance in terms of behavioral correlates would have been a matter for further experimental investigation.

Studies Other than Rorschach Studies

There have been very few studies by psychologists of the personality of epileptics which have not used the Rorschach test. Meyers and Brecher (1941), using the Kent-Rosanoff Word Association Test, found no sig-

nificant differences between a group of idiopathic epileptics and normal persons, matched for age, sex, IQ, and socioeconomic status. Arluck (1941) found no significant differences between the scores of his epileptic and control groups on a level of aspiration test and a personality questionnaire. These results are difficult to evaluate in the absence of evidence that the tests can differentiate significantly between any diagnostic groups whatsoever.

Davies-Eysenck (1950) assessed 38 adults suffering from idiopathic epilepsy, using three tests of neuroticism, which had been shown to discriminate very significantly between normals and neurotics. Even though she selected only those patients who were not mental defectives, were capable of paid employment, and fairly regular in clinic attendance, she found the mean score of the group one standard deviation towards the neurotic end of the scale. There was no correlation between length of illness and degree of neuroticism.

Halstead (1957) compared 28 epileptic children attending ordinary schools, 12 epileptic children attending a physically handicapped school, and 28 attending a special epileptic residential school, with 54 normal children. The investigation was primarily concerned with cognitive abilities and educational attainments, but behavior was assessed from information supplied by parents and schools. Thirty-seven percent of the epileptic group were considered to have good behavior, 35% to have bad behavior (aggressive, destructive, etc.), and 28% to have negative behavior (sullen, timid, oversensitive, etc.). The only variable analyzed which showed a significant correlation with bad behavior was attendance at the special epileptic school. This was, of course, one of the main reasons for referral. Positive but not significant correlations with bad behavior included brain injury, symptomatic epilepsy, frequent seizures, and longer duration of epilepsy. There were no statistically significant correlations between negative or good behavior and other variables, although correlations between good behavior, normal milestones, short duration of epilepsy, having major seizures, and attendance at normal school were all positive. There were no known cases of brain injury among the group with good behavior.

This study is of particular interest because of its representative sample of epileptic children. Unfortunately, the method of making the behavioral ratings is not recorded, and the reliability of the ratings was not assessed.

Gastaut et al. (1955) compared the performance of 60 adult psychomotor epileptics of average intelligence and an unspecified group of idiopathic epileptics on a variety of psychological tests. They described 72% of the psychomotor epileptics as slow and adhesive, with a flat depressed affect which was reflected in their TAT stories. The idiopathic epileptics on the other hand were found to be quick, hyperactive, and emotionally labile. The test results on which these conclusions were based were not reported, and the EEG and clinical criteria by which the epileptic groups were selected were not defined.

Grunberg and Pond (1957) compared from case history records the family background of three groups of children who had attended the Maudsley Hospital: 53 epileptics with conduct disorders, 53 epileptics without conduct disorders, and 33 nonepileptics with conduct disorders. They found very similar adverse

factors (disturbed parental attitudes, marital disharmony, breaks and changes in the environment) were present in the families of both groups of children with conduct disorder, but absent from the background of well-adjusted epileptic children. The method by which the adverse factors were assessed is not described, andthe reliability of the judgments was not assessed. This study confirms the earlier finding of Sullivan and Gahagen (1935) that epileptic children with serious personality or conduct disorders "had almost without exception a poor home environment."

Discussion

The work surveyed above gives no support to the theory that all or most epileptics share a characteristic personality. It is, however, hardly adequate to affirm or refute the other theories that have been advanced. There is some evidence that the incidence of personality disturbance may be high among epileptics, or among some groups of epileptics, and that different types of personality disturbance are associated with different types of epilepsy. From Grunberg and Pond's and Nuffield's studies cited above one may infer that these personality differences result from a complex interaction of environmental and pathophysiological factors. The contribution of heredity has not been studied, although Harvald (1954) has shown that psychosis, psychopathy, suicide, and criminality do not occur more frequently among the relatives of epileptics than in the general population.

Little progress is likely to be made unless future studies define much more precisely than before the characteristics of the epileptic group on whom observations are made. Most of the studies reviewed above have described their patients only in terms of age, sex, and IQ, and classified their epilepsy either as idiopathic or symptomatic, or as involving major or minor seizures. Neither method of classification has proved illuminating, and both obscure factors that there is reason to believe may be important.

Patients showing no signs of brain lesion are said to have idiopathic epilepsy, a condition believed to result from an inherited instability of cerebral function (Lennox, Gibbs, & Gibbs, 1940), and differing markedly from symptomatic epilepsy, which is associated with demonstrable cerebral lesions. However, there is evidence of a multiple etiology in all epilepsies. Heredity has been shown to be an important factor in symptomatic epilepsy (Williams, 1950),while a high incidence of twin births and breech births has been found in the history of patients with petit mal seizures and 3 per second WS EEGs, the classical form of idiopathic epilepsy (Churchill, 1959). The diagnosis of idiopathic epilepsy depends mainly on negative evidence—the lack of evidence of a focal lesion or of a history of brain injury. But as our understanding of epilepsy grows, more precise diagnoses can be made. Many patients, for example, who were classified as idiopathic epileptics in the studies reported above would now be diagnosed as cases of temporal lobe epilepsy, with presumed focal lesions.

One might predict that classification of patients into those having major or minor seizures, or both, would have significance at least in terms of the different psychological effects of these types of seizure on the patient and on society. In fact, however, this prediction is not supported by Halstead's (1957) study. He found a correlation, not significant, between

good behavior and having major seizures. Nor has this method of classification neurophysiological significance. Any type of minor seizure may develop into a major one, and both clinical forms may be associated with focal, diffuse, or no known lesions of cortical or subcortical origin. Hence, in the supposedly homogeneous groups of idiopathic or symptomatic epileptics, or epileptics suffering from major or minor seizures, investigated in the studies reviewed above, it is likely that the most varied range of brain dysfunction and lesion occurred. If personality differences between epileptics are related to neurophysiological factors, these studies could not have revealed them. Unfortunately, the current method of clinical classification of epilepsy is hardly more adequate for research purposes. This can be illustrated with reference to the concepts of temporal lobe epilepsy and petit mal epilepsy. The former term is generally used to denote clinically defined psychomotor seizures occurring in patients with a temporal lobe EEG focus. It has been shown, however, that such a focus can be associated with all kinds of clinical manifestations. Only 46% of one series of epileptics with EEG ictal discharges in the temporal region had clinical psychomotor seizures (Jasper, Pertuiset, & Flanigin, 1951). Moreover, Hill (1949) found that in 110 cases with clinical psychomotor seizures, only 32 had temporal lobe EEG foci. Electroclinical correlations are even lower in children (Glaser & Golub, 1955). Glaser and Dixon (1956) were unable to find any relationship between response to a particular drug, clinical form of seizure, or type of EEG abnormality. Postmortem studies have revealed a great variation in the extent, type, and location of lesions in such patients (Gastaut, 1953).

Thus, while the term temporal lobe epilepsy implies the existence of a condition in which clinical symptoms, EEG findings, pathology and response to drugs are highly correlated, in fact the only common feature among patients so diagnosed might be their common site of discharge in the region of the temporal lobe. Symonds (1954) has suggested substituting the term "the temporal lobe epilepsies" to emphasize the variety of conditions found. For research purposes, it is desirable to select patients by well defined EEG and/or clinical criteria, rather than by diagnostic labels.

The term petit mal epilepsy is also inadequate for research purposes. Originally a clinical term for any minor attack, it took on a more precise meaning when Gibbs, Davis, and Lennox (1935) observed that momentary seizures, with or without falling, are often accompanied by a 3 per second SW EEG pattern. This electroclinical association is often called "true" petit mal, and appears to respond specifically to the dione drugs. Clinically identical momentary seizures may, however, be accompanied by bursts of spikes or by complex wave forms, and at least one third of epileptics with 3 per second SW have no petit mal attacks (Clarke & Knott, 1955; Lundervold, Henriksen, & Fegersten, 1959; Silverman, 1954). Unless both EEG and clinical data are given, the type of minor attack cannot be identified.

A satisfactory classification of the epilepsies awaits greater understanding of their mode of action. Meanwhile, in any attempt to establish the behavioral characteristics of particular groups of epileptics, it is important to select the patients according to a criterion whose reliability is known,

and which there is some evidence to consider significant. Nuffield's study has shown the importance of EEG criteria. The reliability of such criteria has not yet been adequately studied, and the improvement of reliability probably awaits adequate quantification of certain EEG phenomena.

An important variable which has been neglected is the type and amount of medication taken. This is not specified in any study. However, the drugs used to control seizures have important effects on the nervous system, some excitatory, some inhibitory. One might reasonably postulate that personality differences between different groups of epileptics are primarily a function of the differential long term effects of different types of drugs. Loveland, Smith, and Forster (1957) tackled this problem, but their study covered a period of 3 months only, and the numbers in each group were very small.

While it has often been suggested that personality disturbance in epileptics occurs only in the presence of cerebral lesions, this factor has been inadequately studied. Halstead (1957) found a positive but not significant correlation between bad behavior in epileptic children and a history of brain injury, and Vislie and Henricksen (1958) found a tendency for the severity of personality disturbances in epileptic adults to be related to the extent of brain lesion, as evaluated by neurological symptoms, pneumoencephalography, and EEG findings.

Without autopsy it is, however, often difficult to assess the presence, extent, and site of brain damage. A good history of birth injury is sometimes available, but there is no simple relationship between birth trauma and brain injury. In a series of 406 stillbirths and neonatal births, only 24% which had shown signs of cerebral irritation gave postmortem evidence of intracranial trauma (Bound, Butler, & Spector, 1956). Moreover, numerous cases are described in which serious lesions of the brain have been found at postmortem which were not suspected at birth.

Neurological signs may provide positive evidence of brain damage, but large sections of the brain may be damaged without giving rise to neurological signs. Pneumoencephalographic findings are rarely available and can only be regarded as valid criteria of brain damage when the lesion is gross. EEG abnormalities are not in themselves evidence of brain damage, nor do they necessarily represent accurately the locus and extent of brain damage. The EEG focus may be remote from the lesion, and the lesion may be more or less extensive than the focus. This is particularly true of temporal lobe foci, which because of the low convulsive threshold of the temporo-orbital regions may be secondary to a primary lesion elsewhere. Hence, estimates of the extent and locus of brain damage can with the techniques at present available be at best approximate. However, Meyer, Falconer, and Beck (1954) have found temporal lobe lesions in the brains of all epileptics with temporal lobe EEG foci who have come to autopsy, and according to Gastaut (1953) such lesions are more diffuse than those found in patients with any other focal epilepsy. The question arises, therefore, whether the relationship between personality disturbance and temporal lobe EEG foci is a function of the diffuse nature of the lesions rather than their location An investigation of the relationship between these factors would seem to involve correlating behavioral studies and EEG studies with autopsy findings.

The contribution of social class differences to personality differences between epileptics has not been studied, and the social class distribution of different forms of epilepsy is not known. It is often held, however, that temporal lobe epilepsy is more frequent in the lower social classes, and there is some evidence that aggressive conduct disorders are more frequent in children of lower social class (O'Neal & Robins, 1958). It would therefore seem worthwhile to determine what contribution, if any, social class makes to the association between temporal lobe epilepsy and aggressive personality disorders. What has to be studied, in fact, is the behavior of individuals with malfunctioning nervous systems, damaged in different areas at different stages of development, controlled to a greater or lesser extent by different drugs, acting on and responding to different kinds of environment. It is clear that the methodological problems involved are very complex and hardly touched by the usual studies that control age and IQ. While most of the variables cannot be experimentally controlled, some of their interrelationships are open to study.

An additional problem which has received inadequate attention is that of sampling. Epileptics attending a neurological or psychiatric institute, especially over a long period are likely to differ in important respects, including psychological characteristics, from epileptics attending a general hospital. Pond's survey of epilepsy in general practice has shown that many epileptics are attended only by general practitioners, and that those seen as outpatients at hospitals tend to be more often psychologically disturbed (Pond & Bidwell, 1960).

A final problem, not specific to this field, lies in the selection of dimensions of personality and the development of reliable and valid instruments with which to measure them. For many years the only psychological tool used was the Rorschach test, the inadequacies of which have been discussed above. So far the main alternative has been the clinical assessment of behavior. The reliability of this method is generally undetermined, and indeed the development of criteria to assess such behavioral manifestations as hyperkinesis or distractibility is one of the more neglected problems in psychology.

To the present writer, however, the most fruitful approach appears to be one which relates personality to more elementary forms of psychological functioning. In epilepsy it is known that we are studying the behavior of a nervous system which is certainly malfunctioning, often more or less diffusely damaged. The intelligence test scores of epileptics with known or suspected brain lesion tend to be below average (Collins, 1951), but there has not yet been any study of the nature of the psychological functions impaired in these patients. It would seem worthwhile to investigate the relationship between the emotional and behavioral disturbances reported and more general defects, e.g., impairment in discriminative functions, or slowness to condition.

Summary

Five basic theories about the personality of epileptics are outlined, and the extent to which they have been affirmed or refuted by clinical and psychological investigations is considered. The findings of studies which have used the Rorschach test are shown to be contradictory, and the inadequacies of this test for research purposes are pointed out. It is argued that progress in this field

depends on a recognition and study of the complex environmental and pathophysiological factors involved, and on the development of reliable criteria with which to classify epileptics.

REFERENCES

ALAJOUNAINE, T., NEHLIL, J., & HOUDART, R. Influence de la lobectomie temporale sur l'état mental des épileptiques psychomoteurs. *Rev. neurol.*, 1958, **98**, 165–171.

ALTABLE, J. P. Rorschach diagnosis in a group of epileptic children. *Nerv. Child*, 1947, **6**, 22–33.

ARLUCK, E. W. A study of some personality characteristics of epileptics. *Arch. Psychol., N. Y.*, 1941, **37**, No. 263.

BAUGHMAN, E. E. Rorschach scores as a function of examiner difference. *J. proj. Tech.*, 1951, **15**, 243–249.

BAUGHMAN, E. E. A comparative analysis of Rorschach forms with altered stimulus characteristics. *J. proj. Tech.*, 1954, **18**, 151–164.

BINGLEY, T. Mental symptoms in temporal lobe epilepsy and temporal gliomas. *Acta psychiat. neurol. Scand., Kbh.*, 1958, **33**, Suppl. No. 120.

BOUND, J. P., BUTLER, N. R., & SPECTOR, W. G. Classification and causes of perinatal mortality. Part II. *Brit. med. J.*, 1956, **2**, 1260–1264.

BOVET, T. Der Rorschachversuch bei verschiedenen Formen von Epilepsie. *Schweiz. Arch. Psychiat. Neurol.*, 1936, **37**, 156–157.

BRADLEY, C. Behaviour disturbances in epileptic children. *J. Amer. Med. Ass.*, 1951, **146**, 436–441.

BRIDGE, E. M. *Epilepsy and convulsive disorders in children.* New York: McGraw-Hill, 1949.

CAMPBELL, F. A., & FIDDLEMAN, P. B. The effect of examiner status upon Rorschach performance. *J. proj. Tech.*, 1959, **23**, 303–306.

CARP, F. M. Psychological constriction in several projective tests. *J. consult. Psychol.*, 1950, **14**, 268–275.

CHURCHILL, J. A. The relationship of epilepsy to breech delivery. *EEG clin. Neurophysiol.*, 1959, **11**, 1–12.

CLARKE, E. C., & KNOTT, J. R. Paroxysmal wave and spike activity and diagnostic sub-classification. *EEG clin. Neurophysiol.*, 1955, **7**, 161–164.

CLARKE, L. P. Treatment of the epileptic, based on a study of the fundamental makeup. *J. Amer. Med. Ass.*, 1918, **70**, 357–362.

COLLINS, A. L. Epileptic intelligence. *J. consult. Psychol.*, 1951, **15**, 392–399.

DAVIES-EYSENCK, M. Neurotic tendencies in epilepsy. *J. Neurol. Neurosurg. Psychiat.*, 1950, **13**, 237–240.

DELAY, J., PICHOT, P., LEMPERIERE, T., & PERSE, J. Le test de Rorschach dans l'epilepsie. Part III. *Encephale*, 1955, **44**, 46–56.

FALCONER, M. A., HILL, D. N., MEYER, A., & WILSON, J. L. Clinical, radiological and EEG correlations with pathological changes in temporal lobe epilepsy. In M. Baldwin and P. Bailey (Eds.), *Temporal lobe epilepsy.* Springfield, Ill.: Charles C Thomas, 1958. Pp. 396–411.

FIEDLER, M. F., & STONE, I. J. The Rorschachs of selected groups of children in comparison with the published norms. Part II. *J. proj. Tech.*, 1956, **20**, 276–279.

FISKE, D. W. Variability of responses and the stability of scores and interpretations of projective protocols. *J. proj. Tech.*, 1959, **23**, 263–267.

FRIEDMAN, H. F. Comparisons of a group of hebephrenic and catatonic schizophrenics with two groups of normal adults. *J. proj. Tech.*, 1952, **16**, 352–360.

GASTAUT, H. Etude electro-encephalographique. *Rev. oto-neuro-opth.*, 1950, **22**, 301–320.

GASTAUT, H. So-called psychomotor and temporal epilepsy. *Epilepsia*, 1953, **3**, 59–96.

GASTAUT, H., MORIN, G., & LESEVRE, N. Étude du comportement des epileptiques psychomoteurs dans l'intervalle de leurs crises. *Ann. med.-psychol.*, 1955, **113**, 1–27.

GIBBS, F. A. Abnormal electrical activity in the temporal region. In, *The brain and human behaviour.* Baltimore: Williams & Wilkins, 1958. Pp. 278–295.

GIBBS, F. A., DAVIS, H., & LENNOX, W. G. The electro-encephalogram in epilepsy and in conditions of impaired consciousness. *Arch. Neurol. Psychiat.*, 1935, **34**, 1133–1148.

GLASER, G. H., & DIXON, M. A. Psychomotor seizures in childhood. *Neurology*, 1956, **6**, 646–655.

GLASER, G. H., & GOLUB, L. M. The E.E.G. of psychomotor seizures in child-

hood. *EEG clin. Neurophysiol.*, 1955, **7**, 329–339.

GRUNBERG, F., & POND, D. A. Conduct disorders in epileptic children. *J. Neurol. Neurosurg. Psychiat.*, 1957, **20**, 65–68.

HALSTEAD, H. Abilities and behaviour of epileptic children. *J. ment. Sci.*, 1957, **103**, 28–47.

HARROWER-ERIKSON, M. R. Rorschach studies of patients with focal epilepsy. *Arch. Neurol. Psychiat.*, 1940, **43**, 1081–1107.

HARVALD, B. *Heredity in epilepsy.* Copenhagen: Ejnar Munksgaard, 1954.

HILL, D. N. The EEG concept of psychomotor epilepsy. In, *Fourth Congrès Neurologique Internationale*, 1949.

HILL, D. N. Epilepsy. In Lord Cohen (Ed.), *The British encyclopaedia of medical practice: Medical progress.* London: Butterworth, 1957. Pp. 86–99.

HOLTZMAN, W. H. Validation studies of the Rorschach test. *J. clin. Psychol.*, 1950, **6**, 343–351.

JAMES, I. P. Temporal lobectomy for psychomotor epilepsy. *J. ment. Sci.*, 1960, **106**, 543–557.

JASPER, M., PERTUISET, B., & FLANIGIN, H. EEG and cortical electrograms in patients with temporal lobe seizures. *Arch. Neurol. Psychiat.*, 1951, **65**, 272–290.

KAYE, I. What are the evidences of social and psychological maladjustment revealed in a study of seventeen children who have idiopathic petit mal epilepsy? *J. child Psychiat.*, 1951, **2**, 115–159.

KEEHN, J. D. A reinterpretation of the role played by color in the Rorschach test. *Brit. J. med. Psychol.*, 1954, **27**, 89–93.

KELLEY, D. M., & MARGUILES, H. Rorschach case studies in the convulsive states. *Rorschach res. Exch.*, 1940, **4**, 157–190.

KLUVER, H., & BUCY, P. C. Preliminary analysis of functions of the temporal lobes in monkeys. *Arch. Neurol. Psychiat.*, 1939, **42**, 979–1000.

KOGAN, K. L. The personality reaction pattern of children with epilepsy with special reference to the Rorschach method in epilepsy. *Ass. Res. Nerv. Ment. Dis.*, 1947, **26**, 616–630.

LAZARUS, R. S., & OLDFIELD, M. Rorschach responses and the influence of color. *J. Pers.*, 1955, **23**, 356–372.

LENNOX, W. Seizure states. In J. McV. Hunt (Ed.), *Personality and the behaviour disorders.* Vol. 2. New York: Ronald, 1944. Ch. 31.

LENNOX, W., & GIBBS, F. Inheritance of cerebral dysrhythmia and epilepsy. *Arch. Neurol. Psychiat.*, 1940, **44**, 1155–1183.

LIDDELL, D. W. Observations on epileptic automatism in a mental hospital population. *J. ment. Sci.*, 1953, **99**, 732–748.

LISANSKY, E. S. Convulsive disorders and personality. *J. abnorm. soc. Psychol.*, 1948, **43**, 29–37.

LOVELAND, N., SMITH, B., & FORSTER, F. M. Mental and emotional changes in epileptic patients on continuous anticonvulsant medication. *Neurology*, 1957, **7**, 856–865.

LUNDERVOLD, A., HENRIKSEN, G. F., & FEGERSTEN, L. The spike and wave complex: A clinical correlation. *EEG clin. Neurophysiol.*, 1959, **11**, 13–22.

MEYER, A., FALCONER, M. A., & BECK, E. Pathological findings in temporal lobe epilepsy. *J. Neurol. Neurosurg. Psychiat.*, 1954, **17**, 276–285.

MEYERS, R., & BRECHER, S. The so-called epileptic personality as investigated by the Kent-Rosanoff test. *J. abnorm. soc. Psychol.*, 1941, **36**, 413–422.

MINKOWSKA, F. L'epilepsie essentielle, sa psychopathologie et le test de Rorschach. *Ann. med. psychol.*, 1946, **104**, 321–355.

MULDER, D. W., & DALY, D. Psychiatric symptoms associated with lesions of the temporal lobe. *J. Amer. Med. Ass.*, 1952, **150**, 173–176.

NEFF, W. S., & GLASER, N. M. Normative data on the Rorschach. *J. Psychol.*, 1954, **37**, 95–104.

NEFF, W. S., & LIDZ, T. Rorschach patterns of normal subjects of graded intelligence. *J. proj. Tech.*, 1951, **15**, 45–57.

NOTKIN, J. Is there an epileptic personality makeup? *Arch. Neurol. Psychiat.*, 1928, **20**, 799–803.

NUFFIELD, E. J. A. Neurophysiology and behaviour disorders in epileptic children. *J. ment. Sci.*, 1961, **107**, 438–457.

O'NEAL, P., & ROBINS, L. N. The relation of childhood behavior problems to adult psychiatric status. *Amer. J. Psychiat.*, 1958, **114**, 961–969.

PAILLAS, J. E. Aspects cliniques de l'epilepsie temporale. In M. Baldwin and P. Bailey (Eds.), *Temporal lobe epilepsy.* Springfield, Ill.: Charles C Thomas, 1958. Pp. 411–440.

PAILLAS, J., & SUBIRANA, A. Séméiologie neuropsychique: Le lobe temporal. *Rev. oto-neuro-opthal.*, 1950, **20**, 123–218.

PETERMAN, M. G. Behavior in epileptic children. *J. Pediat.*, 1953, **42**, 758–769.

PIOTROWSKI, Z. A. The personality of the epileptic. In P. H. Hoch and R. P. Knight (Eds.), *Epilepsy.* New York: Grune & Stratton, 1947. Pp. 89–109.

POND, D. A. Psychiatric aspects of epilepsy in children. *J. ment. Sci.*, 1952, **98**, 404–410.

POND, D. A. Psychiatric aspects of epileptic and brain-damaged children. *Brit. med. J.*, 1961, **2**, 1377–1382.

POND, D. A., & BIDWELL, B. M. A survey of epilepsy in fourteen general practices: II. Social and psychological aspects. *Epilepsia*, 1960, **1**, 285–299.

PRUYSER, P. W. Psychological testing in epilepsy: II. Personality. *Epilepsia*, 1953, **2**, 23–35.

REIMAN, G. R. The effectiveness of Rorschach elements in the discrimination between neurotic and ambulatory schizophrenic subjects. *J. consult. Psychol.*, 1953, **17**, 25–31.

REVITCH, E. Psychiatric aspects of epilepsy. *J. Med. Soc. N. J.*, 1955, **52**, 634–640.

RICHARDS, T. W. The personality of the convulsive patient in military service. *Psychol. Monogr.*, 1952, **66**(14, Whole No. 346).

ROBERTIELLO, M. Psychomotor epilepsy in children. *Dis. nerv. Sys.*, 1953, **14**, 337–339.

ROGER, A., & DONGIER, M. Correlations electrocliniques chez 50 epileptiques internes. *Rev. Neurol.*, 1950, **83**, 593–596.

RORSCHACH, H. *Psycho diagnostics*. (3rd ed.) Bern: Huber, 1942.

ROSS, W. D. The Rorschach method and clinical diagnosis. *J. ment. Sci.*, 1941, **87**, 331–348.

SHAW, M. C., & CRUIKSHANK, W. M. The Rorschach performance of epileptic children. *J. consult. Psychol.*, 1957, **21**, 422–424.

SILVERMAN, D. Clinical correlates of the spike-wave complex. *EEG clin. Neurophysiol.*, 1954, **6**, 663–669.

STAUDER, K. H. *Konstitution und Wesensanderung der Epileptiker*. Leipzig: Thieme, 1938.

SULLIVAN, E. B., & GAHAGAN, L. On the intelligence of epileptic children. *Genet. psychol. Monogr.*, 1935, **17**, 309.

SYMONDS, C. Classification of the epilepsies with particular reference to psychomotor seizures. *Arch. Neurol. Psychiat.*, 1954, **72**, 631–637.

VISLIE, H., & HENRIKSEN, G. F. Psychic disturbances in epileptics. In A. M. Lorentz de Haas (Ed.), *Lectures on epilepsy*. Elsevier: Van Nostrand, 1958.

WEDEMEYER, B. Rorschach statistics on a group of 136 normal men. *J. Psychol.*, 1954, **37**, 51–58.

WILLIAMS, D. New orientations in epilepsy. *Brit. med. J.*, 1950, **1**, 685–692.

WITTENBORN, J. R., & HOLZBERG, J. D. The Rorschach and descriptive diagnosis. *J. consult. Psychol.*, 1951, **15**, 460–463.

YATES, A. J. The validity of some psychological tests of brain damage. *Psychol. Bull.*, 1954, **51**, 359–379.

ZEHRER, F. A. Investigation of Rorschach factors in children who have convulsive disorders and in those who present problems of adjustment. *Amer. J. Orthopsychiat.*, 1951, **21**, 292–300.

ZIMMERMAN, F. T., BURGEMEISTER, B. B., & PUTNAM, T. J. The intellectual and emotional makeup of the epileptic. *Arch. Neurol. Psychiat.*, 1951, **65**, 545–556.

(Received November 28, 1960)

Psychological Bulletin
1962, Vol. 59, No. 3, 211–223

MEANINGFUL AND UNMEANINGFUL ROTATION OF FACTORS

JOHN W. THOMPSON
University College of North Staffordshire

Whereas hitherto the proponents of simple structure have admitted that a perfectly objective series of rotations which, according to them, would ensure a solution that correctly mirrored reality, still remained beyond their grasp, it now appears to be only a matter of time before electronic computing techniques make undisputably unique solutions possible. In the writer's opinion the fundamental problem as to whether mathematically exact solutions mirror reality will remain, and judgmental methods will not thereby be outmoded. In this article the historical background to problems of rotation is described and the problems discussed. The views of several well-known factor analysts—e.g., Cattell and Burt—on the subject of rotation are compared, and it is suggested that, as there will still be room for judgmental methods factor analysts, in addition to programing exact simple structure, should take up and develop suggestions for rotation made by H. J. Eysenck and W. Stephenson. It is also suggested that insufficient distinction is made between the polarity of factors derived from different kinds of data, and that whereas sometimes it is immaterial whether a factor is bipolar or otherwise, on other occasions it is important, so that in scoring tests involving different kinds of data, investigators ought to bear in mind the polarity of resulting factors. According to whether the factors are unidirectional or bipolar, factor analysts should also pay more attention to the way in which big and small loadings are distributed within individual factors, instead of confining their interest to overall arrangement, as in simple structure, cluster analysis, and proportional profiles. In this way it would help to clear up difficulties connected with polarity which have apparently arisen in the interpretation of Eysenck's T, and would widen the scope of factor analysis as an instrument of meaningful classification. Far from decrying objective rotation, the writer suggests its extension to new applications, provided that a parallel extension is permitted in the development of judgmental methods.

Background to the Present Situation

It is customary, especially among American psychologists, to use transformations of the initial factor matrix rather than to interpret the loadings as derived, but some psychologists maintain that rotation is optional or even undesirable. The question, "Should rotation be entirely a mathematical procedure, or should it be based on 'psychological meaning?' " is historically an old one, and whether or not to rotate, is bound up with (*a*) the status to be accorded to factors—whether they are principles of classification, functional unities, or unique entities, and (*b*) whether factor analysis is to be regarded as a branch of dependent or interdependent statistics—whether its purpose is to prove a hypothesis or to explore. These considerations have been debated at length in a paper by Kendall and Babington Smith (1950), and in articles by Eysenck (1944, 1950,

1952), and considerable space has been devoted to rotational procedures in the main texts on factor analysis, e.g., those of Adcock (1954), Burt (1940), Cattell (1952), French (1953), Holzinger and Harman (1941), Stephenson (1953), Thomson (1939), and Thurstone (1947). In spite of all that has been said, rotation continues to be a source of controversy and ambiguity, Cattell and French maintaining that inadequate rotation is responsible for more failures in factor analysis than anything else. The situation requires clarifying, as the otherwise powerful technique of factor analysis is weakened by widespread disagreement. Moreover it is essential to make such an effort now, for the gap between the psychologists who say that rotation is essential, and those who say it is optional, is likely to widen with the further development of electronic computing machines. We would now appear to be approaching the point at which it will be possible to arrive at a unique simple structure even for correlated factors, a procedure that has previously been too difficult. The writer suggests that objective and judgmental rotation, although rarely employable by the same operator, using the same data, on the same occasion, are nevertheless techniques which are complementary, if a sufficiently broad view is taken of the scope of factor analysis. The writer would accordingly like to see a corresponding forward advance in the development of judgmental methods of rotation.

Eysenck's Use of Rotation

Eysenck has on the whole adopted a middle of the road position. He says (Eysenck, 1944) that whether or not factors are rotated, the various methods of factor analysis, far from being incompatible, are alternative approximations of one another. This point of view is, however, insufficient to account for Eysenck's discovery of the basic attitudes Radicalism-Conservatism and Tendermindedness (R and T), if his claim to have achieved it essentially through a reinterpretation of previous factorial studies by undoing the previous work of rotation (Eysenck, 1954) is to be accepted. Eysenck mentions, in particular, the analyses of Thurstone (1934), Lurie (1937), Duffy and Crissy (1940), and Hatt (1948). But how far, in fact, was the recalculation of centroid factors essential to Eysenck's discovery? Cattell maintains that it is merely a matter of chance whether unrotated factors are meaningful, and on that view it would be just possible, due to chance, for the centroids of several successive analyses to be capable of a common intepretation, but a more probable explanation is that even when the procedure of factor identification is believed to be completely objective, a certain amount of personal judgment is involved, and that in orginally discovering Conservatism-Radicalism and Tendermindedness, Eysenck unconsciously read meaning into the unrotated factors of the several analyses, in support of a brilliant but already formed hypothesis. At any rate, in the course of Eysenck's further investigations, either the chance sequence of meaningful centroids broke down or Eysenck realized that he could not continue to get support for his hypotheses from further unrotated factors, because in the course of analyzing the data in his "social insight" study (Eysenck, 1951) he was forced to admit the necessity for rotation through 47 degrees to equate his centroid factors with the previously identified R and T. Eysenck (1954) has elsewhere demonstrated

that the angle of separation between R and T and Ferguson's (1952) Religionism and Humanitarianism was 45 degrees. From Eysenck's social insight study the equivalent of Ferguson's and not Eysenck's basic attitude factors must therefore have emerged directly, an event which was presumably unexpected, as up to that time Eysenck had cited the direct emergence without rotation of his own factors R and T as a reason for preferring them to Ferguson's Religionism and Humanitariansm. This experience of Eysenck's would seem to dispose of any argument that centroid factors as such have a special meaningful property.

It was Eysenck (1950) who introduced "criterion analysis," using rotation to align a factor with a "criterion column" incorporated in the correlation matrix. This procedure has had a mixed reception. Cattell, for example, has shown a lack of enthusiasm for it, but it is difficult to see the reason for this, as Thurstone himself, although he did not incorporate a criterion column like Eysenck, or provide a mathematical method such as Lubin's (1950) for bringing about the necessary transformation, recommended the inclusion in the factor analytic population of groups having widely varying characteristics. Criterion analysis is quite in accord with the Thurstonian tradition. Cattell's chief objection seems to be that unless criterion groups differ in respect of more factors than one, such groups constitute special populations, but he surely misses the point as the art in using criterion analysis depends on skillfully selecting just such groups which do *not* form special populations.

Cattell's Views on Rotation

Cattell's ideal is Thurstone's simple structure, giving an invariant solution, but he appears to vacillate in his views as to how far a unique solution is at present actually obtainable, especially in the case where factors are correlated. Cattell's (1952) earlier position would appear to allow in practice for more than one solution with approximately simple structure, depending upon the way in which the analyst proceeds in the initial stages of rotation, and the moves he makes when he begins by trial and error to seek out promising positions. More recently Cattell (1957) under the influence of Bargmann (1954) has openly advocated blind rotation, maintaining that a factor analyst should be quite unaware of the nature of the variables with which he is working, and saying that important sharpening of hyperplanes, involving essential changes of position over and beyond refinement, is delayed until the very closing stages. Cattell (1957) quotes Bargmann as saying that several alternative roughly satisfactory "simple structures" can be found which reach the 10% or even the 5% level of significance, but which are nevertheless (*a*) quite spurious in factor terms, and (*b*) emphatically below the significance (often at the .001 level) of the simple structure obtainable after *thorough* exploration of possible rotation postions. It is evident from articles by Cattell (1955), Kaiser (1958), Neuhaus and Wrigley (1954), and Sokal (1958) that analytical solutions are now commonplace. On first thoughts it might appear probable that, as electronic computing techniques continue to improve, alternative solutions will be discovered at even higher levels of significance, so that the mathematically perfect solution might remain forever out of reach—indeed Saunders (1960) seems

to envisage something of the kind. This let-up is, however, only temporary as similar arguments have appeared before. Stephenson, for example, feeling that it was difficult to accept one kind of geometrical substructure as, in principle, the only basis for inference, found consolation in the fact that, in practice, by way of single-plane and other solutions to the rotational problem, Thurstonian procedure allowed far more latitude than might appear at first sight (Stephenson, 1953, p. 41). Second thoughts on the subject indicate that if in allegedly unique solutions there is still room for permissiveness, it is only a matter of time, both in the orthogonal and oblique case, before undisputably unique solutions appear. Even when this occurs, however, it will not decide the issue of objective versus meaningful rotation, for the exact solution need not necessarily mirror reality, and for certain purposes, and particularly with certain kinds of data, it would be advisable to develop also solutions based on judgment. Even Cattell cannot entirely escape the element of judging, for when arguing most strongly in favor of exact solutions, in the next breath he contradicts himself by admitting the necessity for judgmental intervention. Thus he says (Cattell, 1957),

> If, very occasionally, our art of judgment seems to defy a mechanical following of our own rules, it should be remembered that experience necessarily takes in additional considerations beyond the index value, notably; statistically sophisticated estimates of rounding errors; effects of different communality estimates; differences of sample in mean and heterogeneity; degree of reliability of particular tests in particular situations; changes of method in factor extraction, and the degree of "wobble" of the reference vector to be expected in even the best visual simple structure determination. These are matters of good total perspective, and meticulous attention to sampling error and test reliability will avail the experimenter nothing if he shows negligence in evaluating rotational procedures and checking that simple structure has truly been achieved (p. 232).

Cattell also insists that a factor, as given by simple structure, must be in agreement with previous studies and with prevailing opinion. Thus it would seem to be hard to deny the necessity of the investigator's judgment, and the problem when to use it and how to incorporate it satisfactorily in rotation procedure.

Sir Cyril Burt's Position

Burt has tended to get less involved than other factor analysts in controversies on rotation, and it can be argued that in sectioning matrices by visual inspection of correlation clusters he in fact by-passes the problem. Burt (1940, p. 250) maintains that the primary value of factors is descriptive, that they are principles of classification, and that it depends upon the design of the experiment as to whether factors are meaningful or unmeaningful. The previously given quotation of Cattell's is in agreement with the last of these observations. Burt appears to be satisfied with the situation in which the Thurstone school work with correlated oblique first-order factors and arrive at second-order factors, which are orthogonal, while Burt works with orthogonal factors which include basic factors with which he says the second-order factors are approximately equivalent. However, it seems to the writer that Cattell's second-order factors are something more than Burt's basic factors because of Cattell's emphasis on levels. Cattell's second order factors belong to a different "computing realm" from lower-order factors and there is, he says, a risk of confusing the computing

realms by altering the generality of material from one analysis to another, which needs to be specially guarded against. If not, we are told, factors which are ostensibly the same, but actually different, may result. French (1953) describes several analyses with material of differing generality resulting in first-, second-, and allegedly third-order factors, viz., those of Baehr (1952), Lovell (1945), and Thurstone (1947, 1951). Should there be any limit to the number of different factor orders, and if not, do the higher levels mean anything at all? If it is imperative to rotate first-order factors to simple structure, why is it not also necessary to rotate higher-order factors to give a unique solution at that level, and in order that they can be equated with similar factors and not confused with factors at other levels in different analyses? This *reductio ad absurdum* in connection with higher-order factors suggests that the proponents of simple structure have no grounds for a claim of logical self-sufficiency, and for this reason alone investigators should equip themselves with a variety of rotational procedures, some of which are judgmental.

Stephenson's "Simplest Structure"

Stephenson (1953, p. 37) says that what Thurstone describes as simple structure is in fact the counterpart in variance design, of confounded or other complex designs for structured samples (in Thurstone's case for samples of persons), the explanations offered for Thurstone's various factors being imputations placed upon the persons who could be represented in balanced or other factorial designs for samples of persons. Stephenson says that the difficulty in R technique is that only one attribute can be measured at a time, and goes on to say (p. 40) that all scientific behavior is relatively specific to each experimental situation. As mentioned above, according to Stephenson, the virtue of Thurstone's centroid method, far from being uniqueness of simple structure, is its alleged permissiveness, which Stephenson regards as admirably suitable for the doctrine in experimentation of "the concreteness of inferential behavior." Stephenson believed that Thurstone, in so far as he sought to limit this permissiveness, broke the axiom of concreteness of behavior.

Stephenson (1953) describes his attitude to rotation as follows,

> In our own case the rotations we pursue follow two broad principles. For unstructured samples we seek to determine sometimes what orthogonal structure best fits the data, for a balanced block design of effects, usually for two levels each [representing positive and negative loadings, respectively, on the factor]. A balanced block design is called a case of "simplest structure" to distinguish it from Thurstone's concept of "simple structure." Ours are always orthogonal, but attention is also given to some properties of the structure which are widely overlooked in multiple factor analysis. We not only seek to "explain" factors α, β, λ . . . but we also ask that all possible combinations of the factors such as $\alpha\beta$, $\beta\lambda$, $\alpha\beta\lambda$ should be explained. The *interpretative* power of a factor rests in the combination it helps to explain, as distinct from its *analytic* power which conerns the explanation it provides in Thurstone's sense vis-à-vis a primary factor (p. 41). (Copyright 1953 by the University of Chicago)

In a further explanatory passage (*ibid.*, p. 108) Stephenson says that the search for simple structure has led to interest being centered on tests which are "pure" with respect to one factor, that is, which are loaded in that factor and no other, whereas if there are three factors any variable, as far as being significantly loaded is concerned, may be related to the three in any of eight ways, and these ways determine whether the variable

is "pure" or "mixed" with reference to the factors. Whereas, says Stephenson, it has become almost axiomatic in factor analysis to regard the pure cases of central importance, in fact the mixed combinations as well as the pure should be taken into consideration, as also indeed should the null case—the variable not loaded on any of the three factors. Stephenson maintains that only full consideration of all the possibilities, "pure," "mixed," and "null" gives adequate explanation, and the position which gives this full explanation is that of simplest structure. Stephenson rotates with a view to obtaining a small number of factors which, together with their combinations, account for the same data, instead of seeking a solution involving as many pure classes as possible. Probably because it admittedly involves judgement as well as mathematics, Stephenson's approach has received an even less enthusiastic reception than Eysenck's criterion analysis from the Thurstone school.

At this point the following argument for developing judgmental as well as objective methods of rotation is relevant. Although it is usually considered an advantage that analytical methods of rotation such as quartimax (Neuhaus & Wrigley, 1954) diminish the chances of any particular variable being overrepresented or underrepresented in the factor constellation as a whole, for specific purposes such one-sidedness may actually be demanded. Hence the need for judgmental as well as analytical procedures. To take a simple illustration, when a person with a British railway ticket travels from London to Weymouth, the route printed on the ticket is "via Theale," and similarly when traveling from London to Bournemouth it is "via Sway." The more important the route, the less important is the station defining the route, Theale and Sway, in particular, being small places meaning nothing to the ordinary traveler. Such stations are, however, important to the railway company for discriminating between different journeys and form part of a system of meaningful coordinates. In a similar way, it may require a different kind of factor constellation to satisfy an expert with a specific purpose, than to fulfil the requirements of an analyst whose purposes are less exact and more general.

Extension of Criterion Analysis and Simplest Structure

An essential feature of criterion analysis is incorporation within the correlation matrix of a column of correlation coefficients between each of the "tests" and the criterion, the latter being usually a characteristic or behavior index differentiating criterion groups. The criterion can, if desired, be established by means of preliminary experiments involving Eysenck's (1950) principle of double maximization which entails alternate purification of criterion and test. The following extension of the use of the criterion is suggested. In his study of basic attitudes Eysenck incorporated the criterion in the factor matrix only in the main investigation, but omitted it in subsequent studies on smaller populations in specific areas and under varying conditions, taking it for granted that the R questionnaire continued to measure Radicalism-Conservatism. It would be better in some investigations to choose a criterion which could be included in subsequent studies as well as in the initial investigations. Such a criterion would have to be long enduring, concrete, easily identifiable, and in-

dependent of location—which may, at first sight, appear to be asking for too much, but such criteria could be found, and the search would be rewarding. It is suggested that Stephenson's notion of simplest structure, in accordance with which the the factor analyst looks for a few factors that, together with their combinations, account for the data, could be helpful in choosing a criterion, and that such a criterion would be superior to the markers carried for the purpose of obtaining near-unity loadings that are customarily used to identify factors in successive experiments, as such markers are like dummy tracer mixed with live ammunition, there being no guarantee that superficial characteristics which are easy to follow are consistently correlated in successive experiments with the remaining experimental material. A criterion selected in the manner suggested could keep a whole series of investigations in alignment without ambiguity of interpretation.

Polarity in Factor Rotation

It is also proposed that in rotation increased importance should be attached to the polarity of factors; this suggestion can be conveniently illustrated by further reference to criterion analysis—in the case of a bipolar factor, the criterion would, of course, also be bipolar. In criterion analysis certain questions arise regarding the nature of a continuum, for Eysenck has maintained that criterion analysis involved more than the mere rotation to a criterion as had been previously employed, and has claimed that in the case of neuroticism, by incorporating within his matrix of correlations, what he called "the criterion column," it was actually demonstrated that the putative factor of neuroticism formed a quantitative continuum (Eysenck, 1950). Using biserial or tetrachoric correlations, he derived the criterion column by correlating a number of tests that discriminated between normals and neurotics, with what he called "the normal-neurotic dichotomy." These tests had been administered to patients at Dartford by Himmelweit, Desai, and Petrie (1946) and Eysenck subsequently factor analysed the data. His argument regarding the neurotic continuum was as follows—if neuroticism was a continuous variable, the group of normal patients taken alone should include persons differing in neuroticism. Persons in this normal group who were more neurotic, should obtain higher scores on those tests which discriminated more markedly between the criterion groups, and moreover such tests should, on the average, be positively correlated. Also, according to Eysenck, if neuroticism was a continuous variable, and the tests were to be factor analyzed, the test which best discriminated normals from neurotics should have the highest loadings on the factor of neuroticism and vice versa, and the other tests should have factor loadings which were proportional to the criterion values. To test these assumptions, Eysenck then performed a factor analysis, and rotated the first summation or centroid factor into maximum correlation with the criterion column, maintaining that if the hypothesis that neuroticism formed a quantitative continuum was not substantiated by the data, no amount of rotation would succeed in giving any but chance correlation between the criterion column and the corresponding factor, and that a high correlation between the rotated factor and the criterion would constitute visible support for the assumption of a neurotic con-

tinuum. Eysenck did in fact obtain a high correlation, and maintained that this was a convincing demonstration that neuroticism formed a quantitative continuum, at the one extreme of which were to be found hospitalized neurotics, while what he described as "so-called normals" were to be found all the way from "near-neurotic and neurotic" to "the conspicuously nonneurotic, mature, stable, and integrated type of personality."

It would appear from the above account that Eysenck regards neuroticism as unipolar, i.e., measured along what the writer has elsewhere called a unidirectional continuum involving measurement from 0 to unity (or, if it is preferred, infinity) and not along a bipolar continuum from -1 to $+1$ (Thompson, 1961). But although he envisages neuroticism as unipolar, Eysenck evidently regards extraversion as bipolar, for his usual practice is to speak of "extraversion-introversion" (e.g., Eysenck, 1956a). And, although Eysenck always talks of Conservatism-Radicalism he sometimes uses the term Tendermindedness instead of Tendermindedness-Toughmindedness—is this a convenient abbreviation, or does this indicate disregard of important questions appertaining to polarity? The use of the notation R and T certainly bypasses any such problems if they exist. This prompts the writer to ask the following question. "Had Eysenck so wished, would it have been legitimate, in dealing with the normal population only, to have altered the system of scoring of his tests and the signs of his correlations, and thus to have had a bipolar factor of neuroticism?" It would, in the writer's opinion, have been permissible, had Eysenck wanted to do this, *in this particular case.* Using the normal group only, it was immaterial whether all the loadings of the resulting factor of neuroticism were positive, or whether a symmetrically bipolar factor resulted from rotation—this kind of polarity should, the writer has suggested (Thompson, 1961), be called tautologous or "convertible" bipolarity.

Eysenck's factor analysis was, however, performed on the normal group only. Supposing Eysenck had included the hospitalized group of neurotics as well as the normal group of patients in his factor analysis, what would the position have been then? In the writer's opinion, the inclusion of the hospitalized group would have introduced complications, and if there were qualitative differences between the normal and hospitalized groups, a unidirectional factor would no longer suffice, and it would have been absolutely necessary for neuroticism to be represented as a bipolar factor. This is what the writer means by "inconvertible bipolarity." From the point of view of explaining the difference between convertible and inconvertible bipolarity, it is highly convenient that in the Dartford study the criterion groups were different in certain important respects from those used in similar studies reported by Eysenck. Thus the Dartford neurotics in the investigation undertaken by Himmelweit, Desai, and Petrie (1946), and whose data Eysenck subsequently used to demonstrate the existence of the neurotic continuum, were less severely neurotic than the cases used previously, while it was found that the normal group were less stable than expected. According to Himmelweit, Desai, and Petrie (1946),

> Dr. Maxwell Jones, who was in charge of the Dartford patients and had previously worked at Mill Hill, reported that the patients were

less seriously ill. He found that the breakdown of the Dartford patients was to a larger extent determined by exceptional environmental stress and to a lesser extent by neurotic predisposition. On the other hand the normal group was probably less well adjusted than the groups of actual soldiers who had previously served as controls. Hospitalisation tends to bring out neurotic attitudes; it is also likely that some of the surgical complaints were the result of accident proneness which has been shown to be correlated with neuroticism (p. 176).

In the present writer's opinion, if a factor analysis of the Dartford neurotics and normals taken together had been performed, there would still not have been sufficient qualitative differences between the two groups to make it a matter of importance whether the factor of neuroticism was bipolar or whether it was unidirectional (convertible bipolarity). But if, on the other hand, wounded and otherwise healthy soldiers were pooled with severe neurotics like those at Mill Hill in a factor analysis, a bipolar factor would be more than a matter of form; it would be obligatory. This is what is meant by inconvertible bipolarity. The example of the Mill Hill and Dartford groups is not an isolated one; whenever factors are derived from the ordering of preferences, a parallel to the qualitative difference between the healthy soldiers and severe neurotics is to be found; this can be demonstrated in the measurement of values. Brogden (1952) has thus remarked on the qualitative difference between the extremes of bipolar factors measuring values (e.g., "Ideality-Practicality"), observing that subjects who score highly on such factors have characteristics for which there are, properly speaking, no counterparts in those scoring low, and vice versa. There is in an esthetic versus politico-economic factor, such as might be derived from the Allport-Vernon Study of Values, nothing that concerns politics or economics on the esthetic side of the factor; such a factor marks a separation between opposites that are not only on a continuum but are also qualitatively different, and the relationship involved is more complex than in a unidirectional or only nominally bipolar factor which represents, let us say, "frequency of eyeblink."

The distinction between the two kinds of polarity may, it is now suggested, be of particular importance in politics. As far as their party programs are concerned and probably more generally than is realized, politicians of different complexions do not just have different attitudes to the same things; they are largely concerned with incompatible aims and objects, and are not even interested in many of the things which are of considerable concern to their opponents. A comparison of the programs of any two rival political parties should be sufficient to demonstrate this. Unless exceptional circumstances prevail, purely political ideologies are "exclusive," not "collusive," and to represent political conservatism-radicalism as unidirectional would be to assume a coalition where none exists. It is interesting to note in this connection that Cohen (1951) feels strongly that no one, especially a person engaged in politics, should be allowed to state his own case or to criticize his opponent until he has satisfied him and any others present that he has grasped his opponent's point of view (Cohen, 1958, p. 152). Although at first sight this appears an admirable suggestion, there is another side to the picture, and Cohen's suggestion, if carried to excess, might destroy the inconvertible polarity which the present writer considers is an essential

feature of all value judgments, and make it impossible to pass any criticism whatsoever. In any event, as long as politicians eschew role reversal, and behave as they do at present, the fact that political ideologies are exclusive and not collusive must be taken into consideration when factor analyzing political attitudes.

Having called attention to the existence of two kinds of polarity, what practical steps are factor analysts to take to differentiate between them? Menger (1959) has pointed out, in connection with measurement theory, that although mathematical formulae may appear formidable to the unitiated, in actuality mathematicians are hampered by serious deficiencies in respect of notation, and cannot easily express certain essential mathematical notions. There would thus appear to be little hope of looking to mathematicians for an immediate formulation of the difference between convertible and inconvertible polarity—a subject with which logicians are more familiar than mathematicians—as the latter can hardly cope with their existing difficulties. A mathematical formulation should, however, remain the long-term objective. In the meantime, it is suggested one method of distinguishing between the two kinds of polarity in factor analysis might be by way of judgmental rotation, i.e., by treating inconvertibly bipolar factors such as those for values and political attitudes in a different way from other factors. Such difference in treatment would concern the manner in which individual factors were spread over "tests." It is suggested that in rotating inconvertibly bipolar factors the object should be to concentrate the loadings at the extremities and to leave a transition zone in the middle, whereas in other factors (i.e., those in which all the loadings are positive, or in which bipolarity is merely nominal) the aim should be even spacing all along the continuum (to have as full a range of loadings as possible). But quite apart from this question of polarity the writer thinks that more attention should be paid to loadings *within* individual factors, that too much regard is paid to the proportion contributed by factors to the total variance, and that an excessive emphasis is placed upon the general pattern or configuration of factors taken together. However, the proposed suggestions are not intended to have absolute priority, and the size of factor loadings and their spacing are, it is realized, bound to be determined in part by the influence of the remaining factors. Nor is it intended to convey that it is a matter of great importance, if loadings are irregularly dispersed or vary greatly in the case of factors embracing only a few tests. What is suggested, is that when inconvertibly bipolar factors embrace a large number of tests, the concentration of loadings at the extremities of the factor (for which the term "antithetical concentration" is now suggested) may prove decisive in choosing between alternative, and otherwise equally acceptable, rotations, and that what is suggested should be called "even spacing" i.e., the uniform dispersal of loadings over the continuum may prove similarly decisive in the case of factors which are not inconvertibly bipolar.

In respect of antithetical concentration and even spacing, a "ready-reckoner" in the form of a table or nomogram might be of use to factor analysts to indicate whether one rotation was preferable to another. An investigator might want to know, in the case of an inconvertibly bipolar

factor, whether, e.g., loadings of −.9 −.9 −.8 −.7 and +.5 +.7 +.8 +.8 were to be preferred in respect of antithetical concentration to loadings of −.9 −.7 −.4 −.3 and +.9 +.9 +.9 +.9, or whether in the case of a unidirectional factor loadings of .9 .8 .5 .5 .2 were preferable in respect of even spacing to .8 .8 .7 .4 .2. (It would sometimes not be easy to come to a decision on the basis of inspection.) The construction of a ready-reckoner would not necessarily involve difficult computational procedures; the simple averaging of appropriate sums or differences might suffice, the situation resembling in this respect, that in which L. W. Ferguson (1944), in constructing his basic attitude scales for Religionism and Humanitarianism discovered that the rounded-off correlation coefficients between items and the attitude continuum yielded item weights just as satisfactory as those derived from more complicated procedures. But, bearing in mind that only in a superficial sense is a factor loading of .6 twice that of a loading of .3, it might be necessary for the accurate calculation of indices of antithetical concentration and even spacing to construct a nomogram on more elaborate principles, based on the rate at which the coefficients drop off from extreme to lower values.

One reason why it might be worthwhile to construct a nomogram for investigators is that in the past insufficient appreciation of problems of polarity has apparently led to confusion, and in particular, may lie at the root of the misunderstanding between Eysenck and other investigators concerning Eysenck's T, reported in the *Psychological Bulletin* (Christie, 1956a, 1956b; Eysenck, 1956b, 1956c; Rokeach & Hanley, 1956a, 1956b). In the course of this controversy, criticisms were made regarding Eysenck's scoring, and also regarding his contention that his failure, after exhaustive attempts, to find more than a few items with high loadings on T and no other factor, could be construed as support for the view that T was difficult to define and was probably closely related to extraversion-introversion acting in conjunction with R.

An area in which the use of an index of antithetical concentration might have special applications is that of *Q* technique, where factors are defined by clusters of subjects who have, e.g., a similar personality make up. Here it should be possible to compare factor rotations in which the antithetical concentration was maximised with respect to different persons or groups. Given the right kind of data, it would in this way be possible to provide a flexible measure of group interaction, and to narrow the gap between factor analysis and field theory. A further possibility might be the production of asymmetrically bipolar factors with more loadings on one side of a neutral point than on the other corresponding to the bias associated with taking the data relating to a given group as factorially neutral; the same effect could be studied in R technique by taking a particular test or group of tests as a neutral centre of factorial reference. From developments such as these a technique of psychometric relativity could emerge, which, although having a resemblance to field theory, would remain a factorial technique in affinity with Stephenson's doctrine of the "concreteness of inferential behavior."

Conclusions

There is still room both for judgmental and objective methods of

rotation and the two techniques, although separate, should be complementary. In objective methods of rotation, e.g., simple structure, cluster analysis, and proportional profiles, consideration is given to the arrangement of all the factors together, but insufficient regard is paid to the distribution and size of loadings within individual factors. The terms "antithetical concentration" and "even spacing" are suggested for use with inconvertibly bipolar and other factors, respectively. It is hoped in this connection that these concepts will lead to greater discrimination in the analysis of different types of data and that they will prove useful principles for factor rotation. It is also hoped that the distinction between the two kinds of polarity will help to clear up past misunderstandings regarding Eysenck's criterion analysis, a form of factor analysis which could with advantage be adapted and used more extensively. This is also true of rotation in accordance with principles suggested by Stephenson. The writer admits the uses of simple structure, but is also in sympathy with Stephenson's tendency to emphasize the particular, and takes the view that the technique of rotation to be employed should depend upon the nature of the data, as well as on the aims of the investigator. There is thus room both for mathematically exact solutions and for judgmental rotation. With the development of electronic computers the present time is especially favorable for developing exact mathematical solutions, and these will become increasingly available so the popularity of objective methods of factor rotation is thus assured, but care is needed to ensure that judgmental methods are not forgotten and that their potential use in new applications is also recognized.

REFERENCES

ADCOCK, C. J. *Factorial analysis for non-mathematicians.* Melbourne: Univer. Press, 1954.

BAEHR, M. E. A factorial study of temperament. *Psychometrika*, 1952, **17**, 107–126.

BARGMANN, R. *Mitteilüngsblatt für mathematische Statistik.* Sonderdrack Wurzburg: Physica-Verlag, 1954.

BROGDEN, H. E. The primary personal values measured by the Allport-Vernon Test, "A Study of Values." *Psychol. Monogr.*, 1952, **66**(16, Whole No. 348.)

BURT, C. *Factors of the mind.* London: Univer. London Press, 1940.

CATTELL, R. B. *Factor analysis.* New York: Harper, 1952.

CATTELL, R. B. *Personality and motivation structure and measurement.* New York: World Book, 1957.

CATTELL, R. B., & CATTELL, A. K. S. Factor rotation for proportional profiles. Analytical solution and an example. *Brit. J. statist. Psychol.*, 1955, **8**, 83–92.

CHRISTIE, R. Eysenck's treatment of the personality of communists. *Psychol. Bull.*: 1956, **53**, 411–430. (a)

CHRISTIE, R. Some abuses of psychology. *Psychol. Bull.*, 1956, **53**, 439–445. (b)

COHEN, J. The technique of a role reversal. *Occup. Psychol.*, 1951, **25**, 64–66.

COHEN, J. *Humanistic psychology.* London: Allen & Unwin, 1958.

DUFFY, E., & CRISSY, W. J. E. Evaluative attitudes as related to vocational interests and academic achievements. *J. abnorm. soc. Psychol.*, 1940, **35**, 226–245.

EYSENCK, H. J. General social attitudes. *J. soc. Psychol.*, 1944, **19**, 207–227.

EYSENCK, H. J. Criterion analysis: An application of the hypothetico-deductive method to factor analysis. *Psychol. Rev.*, 1950, **57**, 38–53.

EYSENCK, H. J. Primary social attitudes and the social insight test. *Brit. J. Psychol.*, 1951, **42**, 114–122.

EYSENCK, H. J. Uses and abuses of factor analysis. *Appl. Statist.*, 1952, **1**, 45–49.

EYSENCK, H. J. *The psychology of politics.* London: Routledge & Kegan Paul, 1954.

EYSENCK, H. J. The inheritance of extraversion-introversion. *Acta psychol., Amst.*, 1956, **12**, 95–110. (a)

EYSENCK, H. J. The psychology of politics and the personality similarities between fascists and communists. *Psychol. Bull.*, 1956, **53**, 431–438. (b)

EYSENCK, H. J. The psychology of politics: A reply. *Psychol. Bull.*, 1956, **53**, 177–181. (c)

FERGUSON, L. W. A revision of the primary social attitude scales. *J. Psychol.*, 1944, **17**, 229–241.

FERGUSON, L. W. *Personality measurement.* New York: McGraw-Hill, 1952.

FRENCH, J. W. *The description of personality measurements in terms of rotated factors.* Princeton: Educational Testing Service, 1953.

HATT, P. Class and ethnic attitudes. *Amer. sociol. Rev.*, 1948, **13**, 36–43.

HIMMELWEIT, H. T., DESAI, M., & PETRIE, A. An experimental investigation of neuroticism. *J. Pers.*, 1946, **15**, 173–196.

HOLZINGER, K. J., & HARMAN, H. H. *Factor analysis.* Chicago: Univer. Chicago Press, 1941.

KAISER, H. F. The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 1958, **23**, 187–200.

KENDALL, M. C., & BABINGTON SMITH, B. Factor analysis. *J. Roy. Statist. Soc.*, 1950, **12**, 60–94.

LOVELL, C. A study of the factor structure of thirteen personality variables. *Educ. psychol. Measmt.*, 1945, **5**, 335–350.

LUBIN, A. A note on "criterion analysis." *Psychol. Rev.*, 1950, **57**, 54–57.

LURIE, W. A. A study of Spranger's value types by means of factor-analysis. *J. soc. Psychol.*, 1937, **8**, 17–37.

MENGER, K. Mensuration and other mathematical connections of observable material. In C. W. CHURCHMAN AND P. RATOOSH (Eds.), *Measurement: Definitions and theories.* New York: Wiley, 1959.

NEUHAUS, J. O., & WRIGLEY, C. F. The quartimax method: An analytical approach to orthogonal simple structure. *Brit. J. statist. Psychol.*, 1954, **7**, 81–91.

ROKEACH, M., & HANLEY, C. Care and carelessness in psychology. *Psychol. Bull.*, 1956, **53**, 183–186. (a)

ROKEACH, M., & HANLEY, C. Eysenck's tendermindedness dimension: A critique. *Psychol. Bull.*, 1956, **53**, 169–176. (b)

SAUNDERS, D. R. A computer program to find the best-fitting orthogonal factors for a given hypothesis. *Psychometrika*, 1960, **25**, 199–205.

SOKAL, R. R. Thurstone's analytical method for simple structure and a mass modification thereof. *Psychometrika*, 1958, **23**, 237–257.

STEPHENSON, W. *The study of behaviour.* Chicago: Univer. Chicago Press, 1953.

THOMPSON, J. W. The bi-polar and unidirectional measurement of intelligence. *Brit. J. Psychol.*, 1961, **52**, 17–23.

THOMSON, G. *The factorial analysis of human ability.* London: Univer. London Press, 1939.

THURSTONE, L. L. The vectors of mind. *Psychol. Rev.*, 1934, **41**, 1–32.

THURSTONE, L. L. *Multiple factor analysis.* Chicago: Univer. Chicago Press, 1947.

THURSTONE, L. L. The dimensions of temperament: Analysis of Guilford's thirteen personality scores. *Psychometrika*, 1951, **16**, 11–20.

(Received December 27, 1960)

Psychological Bulletin
1962, Vol. 59, No. 3, 224–242

STRATEGIC, METHOD, AND STYLISTIC VARIANCE IN THE MMPI[1]

JERRY S. WIGGINS
Stanford University

Stylistic consistencies in subjects' responses to personality tests (Jackson & Messick, 1958) are mediated by the form in which the stimulus materials, or items, are presented to the subject. The limitations on response alternatives imposed by the test format do not *create* stylistic consistencies in subjects—rather they define the sample of stylistic responses to be measured on any occasion. Nevertheless, there are some response measures that are so closely related to the structure of the instrument that they appear to tell us more about the measurement method than they do about the subjects. Consequently, investigations of response styles may lead to the detection of stylistic consistencies which are personality traits in their own right or they may uncover sources of method variance that increase our understanding of the instruments themselves. An additional consideration with psychiatric inventories, such as the MMPI, is the assessment strategy of criterion group comparison. This particular scale construction design employs a statistical definition of personality dimensions which has profound implications for the kinds of styles that may be sampled with the instrument.

The present paper draws a distinction among three sources of variance that contribute to assessment with the MMPI. These will be termed: strategic, method, and stylistic variance, respectively. By *strategic* variance is meant variation in test scores that may be attributed to the overall strategy of constructing scales to discriminate between criterion groups and a normative population. The major dimension of variance here is variance in relative *communality* between any given subject and a predetermined normative group. For selection purposes, it is convenient to find dimensions of noncommunality between criterion group members and the normative group such that any individual may be classified as being more or less like the criterion group or the normative group on this particular predictive dimension.

The particular souce of *method* variance at issue with inventories such as the MMPI is the employment of "true-false" or "agree-disagree" item response options which act as constraints on assessment through communality concepts. That is, although the statistics of criterion group discrimination demand a strategy for estimating individuals' communality on a given dimension, the method whereby this is accomplished with the MMPI is one in which an individual may indicate communality or noncommunality by answering "true" or "false" as the case is determined by item phrasing. Although this limitation on item response may permit assessment of general stylistic tendencies to "agree" or "disagree," which may themselves be of predictive significance, we are referring here to the idiosyncratic nature of the particular item pool em-

[1] Portions of this paper were read at the Conference on Personality Measurement, October 13, 1960, Educational Testing Service, Princeton, New Jersey.

ployed with a given instrument. This would include such characteristics of the item pool as the proportion of true and false keyings that exists among the items and the distribution of relative communality or "item popularity" values within the particular item pool.

By *stylistic* variance is meant certain characteristic response consistencies on the part of subjects which may be shown to exist relatively independently of the test itself, but whose detection is to a large extent limited by the appropriateness of a given test as a stimulus for these consistencies to emerge. The specific stylistic dimension considered here is that of "acquiescence versus cautiousness." This unidimensional response style is here given a more narrow definition than is usually the case and an attempt is made to clearly differentiate it from both strategic and method sources of variation. It will be argued that although this stylistic tendency may be considered as important on other grounds, its operation within the context of the MMPI has generally been obscured by the operations of both strategic and method variance in determining test scores.

Item Communality and the Strategy of Criterion Group Assessment

Items that display strong modal tendencies for one of the response options in a normative group have a high degree of *communality*. That is, the normative group is relatively uniform in its endorsement or rejection of the item. Because of the widespread preference among constructors of empirical inventories for the simplest method of item weighting (1 or 0), this implicit "intensity" parameter is seldom systematically taken into account (Guilford, 1954). Nevertheless, it is recognized that very high degrees of communality impose limits on the potential usefulness of an item in the strategy of criterion group comparison (Berg, 1961; Sechrest & Jackson, 1960). Similarly, those who favor the utilization of so-called "subtle-0" items (Meehl, 1945) would recognize that high communality would impose an upper limit since this type of item requires that the criterion group communality exceed that of the normative group. More important, perhaps, is the possible psychological (trait) significance of an "unpopular" or noncommunality response to an item of high communality as opposed to an unpopular response to an item of medium communality. Within the MMPI this is best illustrated by the *F* scale which is, in essence, a noncommunality scale consisting of items of very high communality value keyed in the direction of unpopularity. Answering many more than 4 or 5 of these 64 items in the keyed direction might seriously jeopardize a respondent's outpatient status (Dahlstrom & Welsh, 1960).

It has become customary to refer to items of intermediate or medium communality values as items of high "controversiality" (Fricke, 1957; Hanley, 1957). These are items that are endorsed by about half of a normative group and answered false by the other half. The limits of this range have been generally set at between 40 and 60% endorsement (Fricke, 1957). Although this concept is defined by reference to a pattern of group responses, the term is used as if individual response tendencies were involved. Thus, when these items are referred to as "indifferent," the implication is that individuals within the group are indifferent to that item, rather than that there are two distinct opinions

held by two groups of equal size. This usage seems justified, within limits, to the extent that items of high controversiality tend to be unstable on test-retest (Fiske & Rice, 1955) and that they tend to be rated as being of neutral social desirability value (Hanley, 1957). What has not been emphasized is the fact that the concept of controversiality is used with reference to the responses of normative groups. Items which are controversial for one group may not be so for another and, indeed, if this were not so, such items would make no contribution to the discrimination of nonnormal groups.

As has been observed by several writers (Berg, 1961; Wiggins & Rumrill, 1959) the scaling procedures used by Edwards (1953) to obtain rated "social desirability" values for inventory items may be considered as indirect assessments of item communality values. In this procedure, judges are given a pool of items and a set of instructions that requires them to estimate the degree of desirability of an endorsement of the item by others (Edwards, 1957). Studies to date have employed five-point (Heineman, 1952 unpublished), seven-point (Cowen & Stiller, 1959; Wiggins & Rumrill, 1959), and nine-point (Edwards, 1957; Hanley, 1956) interval scales. These ratings are obtained independently of the communality value of the items, although, as one might expect, substantial correlations have been reported between endorsement frequencies and rated social desirability scale values in certain item pools (Edwards, 1957). Although the prediction of communality values from social desirability ratings has been interpreted as having profound and somewhat sinister implications for inventory assessment (Edwards, 1957), the equally plausible prediction of social desirability ratings from communality values is here considered to stem naturally from the function of item communality in criterion group strategy.

The use of psychiatric criterion groups in the development of the MMPI clinical scales reflected the intent of the test authors (Hathaway & McKinley, 1951) to develop indices of the extent to which a given individual could be said to possess traits or test taking attitudes, or both, which are more typical of a psychiatric group than they are of the population at large. The middle range of such scales may be thought of as reflecting the degree of communality an individual possesses with the Minnesota normals on whatever trait or set of circumstances is involved in the scale. This middle range of the MMPI clinical scales has never been seriously proposed as a definitive criterion of "normality" nor have very fine discriminations been expected of it. We would expect a certain amount of "conformity" to be operative here in the statistical sense that the majority of normals will endorse what is considered to be the acceptable response by the majority of normals. This tendency toward conformity becomes of special interest when it appears in the exaggerated form of "hypercommunality." That is, when the individual's responses become determined, almost exclusively, by efforts to match anticipated cultural patterns of acceptability, rather than by honest assessment of his own position on the trait continuum.

We are speaking then of the three broad categories into which an individual may be classified on the basis of his score on a single criterion-derived scale: he may score high on the scale, in the middle range of the scale, or at the lower end. Individuals scoring high on the scale exhibit a noncommunality with the normative group and by implication are more like the criterion group on whatever

trait is measured by the scale. In terms of Berg's Deviation Hypothesis (Berg, 1955, 1957, 1961) these individuals may be deviant in both a "relative" sense (in terms of differing from normative modal tendencies) and in an "absolute" sense (giving responses typical of a group defined to be deviant on other grounds). Individuals who attain scores in the middle range of the scale may be thought of as being more similar to the normative group than to the criterion group on the variable in question. Since the normative group is usually defined by default, with respect to psychiatric variables, there is considerably less information conveyed by such a middle-range score.

It is a peculiar property of criterion derived scales that individuals scoring very low resemble the normative group but not any of its individual members. The possible significance of this category of scorers is recognized by Sechrest and Jackson (1960) who speak of the "deviantly nondeviant"—i.e., scorers who adhere to the modal responses more consistently than do members of the normative group. The present concept of hypercommunality refers to this same category of low scorers and, in addition, calls attention to the fact that it would logically include those individuals whose responses are chiefly determined by considerations of social desirability or exaggerated attempts to respond in the direction of highest estimated item communality.

Communality and Social Desirability

During the past few years, a number of individual difference measures of the response style of social desirability have appeared in the literature which have, explicitly or implicitly, employed the concepts of controversiality, communality, and rated social desirability (see Table 1). Hanley (1961) has recently suggested a classification scheme for social desirability scales based on whether response frequencies played a role in item selection and whether the social desirability of the items was determined explicitly or implicitly. Another basis for classifying social desirability scales, which has been stressed by the present writer (Wiggins, 1959), is whether the response frequencies of the items were contrasted with those of another group. In the method of *contrasted groups*, the responses of a group of subjects, who because of special instructions or special circumstances are considered to be a group of high social desirability respondents, are compared with the responses of a control group.

The method of contrasted groups was employed by both Cofer, Chance, and Judson (1949) and the present writer (Wiggins, 1959) in the development of two role playing scales of social desirability. The responses of a control group were contrasted with those of subjects instructed to answer the MMPI in terms of social desirability. In the construction of the *L* scale of the MMPI (Meehl & Hathaway, 1946), a group of two clinicians constructed items in terms of social desirability and guessed at the frequencies that would occur in a control group. Two groups were contrasted in spirit, if not in practice. The *K* scale of the MMPI (Meehl & Hathaway, 1946) was constructed by comparing the responses of one group of presumably nonfaking patients with those of another group thought to include a large number of social desirability respondents.[2]

[2] Unfortunately, the interpretation of the *K* scale is further complicated by the fact that an additional set of items was added to the scale that had been shown *not* to discriminate between role playing and control college groups.

TABLE 1

SOCIAL DESIRABILITY SCALES

SD: Ten judges were instructed to answer 149 items from *L*, *F*, *K*, and *MAS* scales in such a way as to give the most socially desirable picture of themselves. Unanimous agreement on 79 items which were reduced to 39 items by item analysis. (Edwards, 1957)

Sd-A, *Sd-R:* Items from Welsh's (1956) Factor Scales *A* (39 items) and *R* (40 items) were rated for both "true" and "false" responses on a 7-point scale by a total of 181 judges. *Sd-A* is a pool of low rated items and *Sd-R* is a pool of moderate items. (Wiggins & Rumrill, 1959)

K: Twenty-two items which differentiated patients with high *L* scores and normal MMPI profiles from a comparable group of patients with abnormal profiles plus eight items which remained unchanged under role playing instructions in normals and also differentiated severe disturbance from normality. (Meehl & Hathaway, 1946)

Ex: Fifty-three items of high controversiality (Hathaway norms) were rated by 92 judges on a 9-point scale of social desirability. Twenty-six high and low rated items constitute scale. (Hanley, 1957)

Cof: Thirty-four items which were unchanged under fake-bad instructions but changed under fake-good instructions in a role playing study utilizing 81 subjects. (Cofer et al., 1949)

Sd: Records of 178 social desirability role players were contrasted with 140 controls to yield 40 differentiating items. (Wiggins, 1959)

L: Two judges made up 15 socially undesirable items that they felt would be frequently endorsed by normals. (Meehl & Hathaway, 1946)

Sx: Eight items were eliminated from *Ex* (see above) to make a balanced scale of 9 "true" and 9 "false" items. (Hanley, 1957)

TSD: Using Heineman's data (1952) on 5-point favorability ratings for all MMPI items by 108 subjects, all items less than 2.5 and greater than 3.5 were keyed for favorability. (Wiggins, 1961)

ESD: Using Heineman's data (1952), the 39 items of "extreme" favorability value (less than 1.5 or greater than 4.5) were keyed for social desirability. (Wiggins, 1961)

Hanley (1957) has developed a scale with reference to two groups that were not directly contrasted with one another. Item controversiality was determined from the Minnesota college norms (Dahlstrom & Welsh, 1960) and controversial items were given to another group to determine their rated social desirability values. Items of high and low rated value were retained and, on rational grounds, it might be predicted that such items would be answered in the keyed direction more frequently by social desirability respondents than by an honest group. In the social desirability scale developed by Edwards (1957), the response frequencies of a single group were sufficient to determine item inclusion. The unanimous agreement of 10 role playing judges determined the direction of item keying without reference to a control group. Similarly, the social desirability scales constructed from judges' ratings of the desirability of items from Welsh's *A* and *R* scales (Wiggins & Rumrill, 1959) were based on a single group of judges.

Table 2 presents data on four characteristics of social desirability scales that are relevant to our discussion of communality and rated social desirability. These characteristics are: communality of items in the scale, rated social desirability of items in the scale, endorsement versus rated social desirability of items in the scale, and the success achieved by the scale in identifying subjects instructed to answer the MMPI in terms of social desirability.

Scale communality is defined as the average proportion of subjects, in a normative group, who answered the items in the direction in which the scale is keyed (social desirability). The communality values in Row 1 of Table 2 were computed from item frequency counts of the records of 140 Stanford students (55 men, 85 women). The social desirability scales have been ordered, in terms of their

TABLE 2

CHARACTERISTICS OF SOCIAL DESIRABILITY SCALES

	SD	*Sd-A*	*Sd-R*	*K*	*Ex*	*Cof*	*Sd*	*L*
Communality								
mean	.79	.66	.55	.53	.50	.35	.30	.20
σ[a]	(.13)	(.18)	(.24)	(.22)	(.19)	(.17)	(.16)	(.16)
Social desirability								
mean	5.70	5.47	5.20	4.65	5.23	4.69	4.72	4.39
σ[b]	(.50)	(.65)	(.24)	(1.20)	(.62)	(.66)	(.74)	(.84)
Endorsement versus social desirability								
r[c]	.91	.55	.56	.46	.05	−.33	−.17	.58
Screening efficiency								
phi[d]	.330	.386	.395	.217	.461	.619	.721	.539

[a] One-hundred and forty college subjects.
[b] Fifty college raters.
[c] Group a versus Group b.
[d] (Wiggins, 1959).

average communality values, from high (on the left) to low (on the right). In the bottom row of the same table are the phi-coefficients reported in a study of the predictive efficiency of social desirability scales in separating a group of 250 social desirability role players from a group of 190 controls (Wiggins, 1959). It can be seen that the average communality value of a scale places a definite limitation on the extent to which it is sensitive to the communality shifts produced by role playing instructions. Thus, with a scale such as Edwards' *SD*, in which the average scale communality is close to 80%, even slight shifts in the direction of social desirability would approach the limit of maximum communality. The social desirability scales with lower average communality values, on the other hand, are potentially sensitive to a much wider range of communality increases and this, among other things, is reflected in the relative success of these scales in identifying social desirability role players.

The average scale social desirability values given in the second row represent the averaged median ratings for all items in the scale when answered in the keyed direction (social desirability). Ratings of the item pool that contained all the social desirability scales were obtained from 50 college students (24 men, 26 women) who rated the social desirability of a true answer on a seven-point scale. Since the ratings were for the desirability of a true answer, it was necessary to reflect the ratings for items keyed false in the social desirability scales.[3] When these median ratings and reflected median ratings are averaged across a given scale, an index of the intensity of a social desirability scale is provided. A close correspondence can be seen to exist between the communality values and the social desirability

[3] This reflection must be considered as only an approximation to the rated social desirability of answering false since previous research (Wiggins & Rumrill, 1959) suggests that empirical values will often differ from reflected values.

values which range from "moderately desirable" in *SD* to "neither desirable nor undesirable" in the *L* scale. Despite this apparent variation with communality values, the actual range of social desirability values is quite small, all scales being within 1.31 points of each other. The communality values, on the other hand, cover a range of 59 percentage points.

The third row presents the correlation between the rated social desirability of a true answer and the number of subjects who actually endorsed the item in an independent group of 140 college men and women. This was calculated in the now standard method for estimating the extent to which social desirability considerations influence responses to items in a given scale (Edwards, 1957). It can be seen that despite the limited variability in average social desirability value, there is considerable variation among the scales in the relationship between endorsement and rated desirability. This ranges from a near-zero correlation in Hanley's *Ex* to a .91 in Edwards' scale. It can also be seen that as the communalities approach controversiality, there is a corresponding decrease in the endorsement-favorability relationship and that as the communalities drop below 50%, the relationship becomes negative—with the notable exception of the *L* scale. It should be recalled that the rationale behind Hanley's *Ex* scale was such that he predicted a near-zero correlation would exist in an "honest" population (Hanley, 1957). To the extent that this population of students is honest, we must conclude that the pool of high communality items that constitute Edwards' *SD* would tend to make us suspicious of almost everyone.

All this, of course, raises the general issue of what properly constitutes an individual difference measure of social desirability and what would be the expected behavior of such a measure under role playing instructions? Edwards is quite explicit about this in his monograph on the social desirability variable (1957). It is his contention that since social desirability is a relatively all-pervasive influence in determining inventory responses, the majority of normals should be considered as possessing a substantial amount of the trait to begin with. Under social desirability instructions, he reasons, scores of the minority who are usually uninfluenced by social desirability will shift in the direction of the majority whose faking scores remain unchanged, and all scores become more homogeneous (Edwards, 1957, pp. 55–56).

In the role playing study already mentioned (Wiggins, 1959), the dissimulation measures (including that of Edwards) exhibited increased, rather than decreased variability under special instructions. More important than this, however, is the actual distribution of scales under the two sets of instructions. Figure 1 presents these distributions for two scales of markedly different effectiveness in differentiating the two instructional groups. The upper part of the figure shows the distribution of the writer's *Sd* under the two conditions. This scale, of course, was more or less "custom-made" for this type of discrimination and it is not surprising that satisfactory separation of the two groups is achieved. Under standard instructions, the *Sd* scale is seen to be relatively normally distributed at the lower end of the possible range of scores. This reflects, among other things, the low communality values of the items in the scale. Under fake instructions, it can also be seen that the faking group takes full advantage of the room for

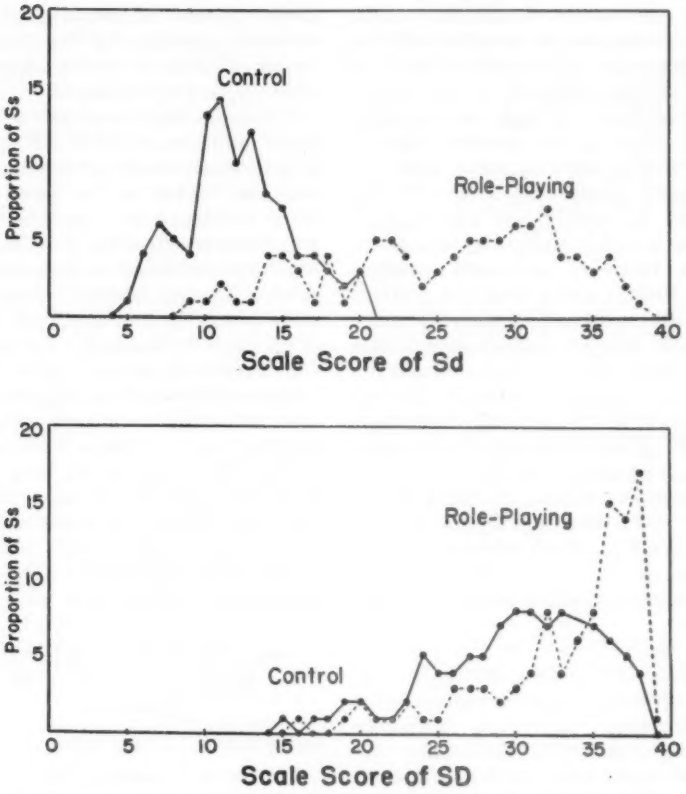


FIG. 1. Distribution of two social desirability scales in role playing and standard instruction groups (from Wiggins, 1959).

shifting in the direction of social desirability by distributing itself along the entire length of the scale—with a median score at a comfortable distance from that of the control group.

The *SD* scale of Edwards, at the bottom of the figure, presents a somewhat different picture. The control group does, as Edwards predicts, show a marked piling up at the upper end of the scale and this would be anticipated by the high average communality value of the scale itself. Under role playing instructions, shifts in the direction of hypercommunality appear to be limited by the already high communality value of the scale.

It should be emphasized that the role playing experiment is not being espoused as the ultimate criterion for social desirability scales. It is readily conceded that there is a certain amount of artificiality to such a procedure that limits its generalization (Hanley, 1961). However, if social desirability is conceived of as a tendency toward communality which at its extreme end exceeds the com-

munality of the average group member, then social desirability scales must themselves be of sufficiently low communality value to allow for detection of this extreme. If we take a pool of items of high communality value scored in the direction of social desirability, and administer them to a standard group and a role playing group, we would not expect large shifts to occur under the two conditions. However, we should not interpret this to mean that the instructions did not produce the style when there is the alternative explanation that the particular item pool did not provide sufficient occasion for the detection of the style in operation. Finally, if social desirability response style is considered to be a style that virtually everyone possesses to a relatively invariant high degree, then it is difficult to see its relevance to the study of individual differences in general or stylistic tendencies in particular.

A recent paper by Crowne and Marlowe (1960) presents an alternative, although parallel, rationale for constructing individual difference measures of social desirability response style. These authors stress the improbability of college students possessing the deviant symptomatology described in Edwards' *SD* items of extreme communality values and argue that "it cannot be determined whether these responses are attributable to social desirability or to a genuine absence of such symptoms" (p. 349). For this reason, the authors selected items "defined by behaviors which are culturally sanctioned and approved but which are improbable of occurrence" (p. 350). No indication is given of the criteria which guided this rational selection although the authors felt they were "avoiding the ambiguities of the statistical deviance approach" by their procedure. Evidence of their success in avoiding items of pathological implications is provided by the mean item rating of 2.9 on a 5-point adjustment scale applied by 10 judges.

That the Marlowe-Crowne Social Desirability scale (M-C *SD* scale) is of sufficiently moderate average communality value to be sensitive to shifts produced by role playing instructions is suggested by its reported normal distribution, in a standard instruction group, around a mean scale score involving less than half (42%) of the items in the scale. The authors contrast this distribution with that of Edwards' *SD* which has negative skew around a mean scale score involving, on the average, some 82% of the items. Since only a fraction of the M-C *SD* scale items were selected from the MMPI, its effectiveness in identifying groups of high social desirability respondents cannot be evaluated from existing data.

Agreement-Disagreement as Method and Stylistic Variance

Definitions of Acquiescence

The term "acquiescence" is used throughout this paper to indicate a general tendency on the part of subjects to agree with a statement or test item when no issue seems at stake. This is, perhaps, a deep-seated cultural bias to be agreeable when it does not cost anything, to "go along with the gang" or the printed word, when no personal inconvenience is attendant upon such acquiescence. At the other end of this postulated dimension is found the tendency to be reluctant to commit oneself to relatively neutral assertions—a disposition to be "cautious" with respect to apparently innocuous issues.

When this dimension is measured within the realm of personality in-

ventories, it is important to note that it refers to response option preferences with respect to a special class of items—i.e., items of "high controversiality" (medium communality) values. Thus, the "acquiescer" will be identified by the number of times he selects true as a response option from a pool of high controversiality items and the cautious individual by the number of false responses to the same item pool. Attempts to assemble pools of high controversiality items from the MMPI by Fricke (1957), Hanley (1957), Fulkerson (1958), and the present writer (Wiggins, 1961) would lead to the conservative estimate that less than 10% of the total pool of MMPI items can be so classified. In addition, as Hanley (1961) has pointed out, even items which meet this statistical criterion of "indifference" may harbor subtle communality variance which may be detected by applying social desirability rating procedures to medium communality items. This is not to deemphasize the possible contribution of acquiescent variance to the MMPI but rather to indicate that highly refined scaling procedures are essential to its *unambiguous* measurement. By unambiguous is meant the separation of acquiescent response style from both communality variance arising from assessment strategy and method variance peculiar to any given item pool.

Since it is difficult to define acquiescence on other than statistical grounds, contrasted criterion groups have not been employed in the development of individual difference measures of this variable (see Table 3). Unfortunately, agreement on the statistical characteristics of acquiescence measures does not appear to be unanimous. Most authors would agree, in principle, that an "acquiescence scale" should be constructed from a heterogeneous pool of items of uniformly high "ambiguity" which are all keyed in the true direction. The statistical criterion of ambiguity (controversiality) has already been discussed. In an ideal normative sample, we would require that the expected communality value of each item be .50 and that the expected scale score be $n/2$, where n = the number of items in the scale.

Fricke (1957) selected items of high controversiality from Hathaway's normative data (Dahlstrom & Welsh, 1960) and keyed them all true as a measure of acquiescence—the *B* scale. Hanley (1961) obtained social desirability ratings on Fricke's *B* scale items and found an imbalance of socially undesirable items to exist.

TABLE 3

ACQUIESCENCE SCALES

B: Using Hathaway's data on endorsement frequencies for a combined college and normal group, 63 items of high controversiality (40–60%) were keyed "true." (Fricke, 1957)

Bn: Using groups of raters totaling 151, Hanley (1961) obtained 9-point desirability ratings on the 63 items from *B* above. The 32 items judged to be of "neutral" value (4–6) were keyed "true." (Wiggins, 1961)

Rb: Eighty-four items of high controversiality (40–60%) in a group of 190 college subjects were rated on a 7-point scale by a different group of 50 subjects. The 27 items of "neutral" value (3.5–4.5) were keyed "true." (Wiggins, 1961)

Acq: From a group of 472 aviation cadets who had been rated for adjustment, 46 items of 40–60% controversiality were found which did *not* discriminate between adjusted and nonadjusted subjects. This pool was reduced to 24 items by internal consistency and they were keyed "true." (Fulkerson, 1958)

At: Keying was changed on half the items in *Sx* (see Table 1) to yield an 18-item "all true" key. (Hanley, 1957)

AT: The number of items, out of 566, that are answered "true" by a subject taking the MMPI.

The 32 items that were judged to be of neutral social desirability by Hanley's raters constitute a revised *B* scale (*Bn*) of neutral social desirability value. Similarly, the writer's *Rb* scale contains items of high controversiality selected from more complete normative data and independently rated for the purpose of retaining items of rated neutral social desirability. Fulkerson's (1958) Acquiescence scale (*Acq*) contains items of high controversiality that have been demonstrated *not* to discriminate between deviant (maladjusted) and normative groups.

The criterion for the "heterogeneity" of an acquiescent item pool is not as clear cut. This may be because of the implicit notion of item "content" which is itself replete with measurement difficulties (Berg, 1959). Cohn (1953) describes an MMPI acquiescent item pool that was "submitted to two clinicians who were in agreement that no particular personality syndrome was being tapped by the item content" (p. 335). Couch and Keniston (1960) using another inventory chose items from "a large number of heterogeneous scales [so that] . . . a total score over all items makes no psychological sense" (pp. 152–53). In developing an acquiescence measure with achievement items, Gage, Leavitt, and Stone (1957) employed items that were "sufficiently difficult and obscure to elicit approximately a 50–50 split of "true" and "false" responses" (p. 98).

The foregoing descriptions of acquiescence item pools imply that the items are not measuring common content dimensions and that consistent agreement is therefore interpretable as acquiescence rather than as a content dimension. A literal interpretation of this implication would be that an acquiescence item pool is lacking in statistical homogeneity—i.e., has low interitem correlations. This line of reasoning has been presented by Hanley (1957) in a slightly different context—that of the rationale for a measure of test taking defensiveness. Applying Hanley's argument to acquiescence measures, we would expect a pure measure of acquiescence to have zero internal consistency in an ideally nonacquiescent population and substantial internal consistency in a population composed of only acquiescent individuals.

Cohn (1953) does not report the internal consistency of his MMPI "Plus scale" so that its lack of content homogeneity must rest on the opinions of his two raters. In this connection, Gage et al. (1957) report that when the items from Cohn's Plus scale were given to four of their judges, there was substantial independent agreement that almost all of the items measured a "tendency to self-disparagement" when answered true which would suggest internal consistency on a content basis.

Couch and Keniston (1960) seem to place a premium on the high internal consistency of their measure of acquiescence: "The first sign of the importance of this measure was indicated by the high (+.85) Spearman-Brown split-half (even-odd) reliability of the 360-item scale. The OAS thus provides a reliable measure of agreeing response set as defined in this study" (p. 153). Assuming, as Couch and Keniston do, that "a total score over all items makes no psychological sense" (p. 153), the substantial internal consistency would suggest that their sample of college men were a highly acquiescent group. However, the distribution of scale scores (which were expressed on a seven-point Likert scale) was from 3.1 to 4.5 which means that *none* of the

subject's score fell in the "Agreement" end of the scale (i.e., greater than 4.5). Moreover, although the overall mean score of the 61 subjects was 3.9, less than 1% of all responses fell in the "No Answer" (4.0) category and the response data suggest that the majority of subjects gave equal proportions of agreeing and disagreeing answers.

Gage et al. (1957) are even more explicit in their insistence upon an internally consistent acquiescence scale. Their group of 50 "difficult" information items scored all true, had a difficulty level of approximately .50, and a corrected split-half reliability of .68. Another group of 40 "easy" information items, scored all true, had a difficulty level of approximately .75, and a corrected split-half reliability of .09. The authors conclude: "That difficult and ambiguous items are required to elicit the acquiescence set is again demonstrated by the fact that the reliability of the acquiescence score on the easy information items was only .09" (p. 98). The subjects were 118 graduate students in education who were apparently not thought, on a priori gounds, to be a particularly acquiescent group.

It should be apparent that the measurement of acquiescence provides a psychometric paradox which has only been partially recognized. To the extent that acquiescence may be said to contribute variance to a given measure (be it criterion-valid or systematic error variance) it must do so with some acceptable degree of reliability. The *internal consistency* and generality of the tendency to agree have been stressed from the beginning (Cronbach, 1942, 1946, 1950). Nevertheless, the agreement tendency itself has been defined with respect to *heterogeneous* and ambiguous items in such a manner that internal consistency in a presumed measure of acquiescence may itself be interpreted as content variance which would vitiate the acquiescence measure on logical grounds. It has been proposed that Hanley's (1957) rationale for constructing a measure of test taking defensiveness may be applied to acquiescence measures as a partial resolution of this apparent paradox. Here, as with social desirability, the importance of specifying theoretical expectations concerning the amount of the style to be found in a given group is apparent.

Table 4 presents data on the statistical characteristics of the MMPI acquiescent item pools already discussed. These calculations are based on the records of 100 college subjects (50 men, 50 women) who took the MMPI as a supplemental requirement of Introductory Psychology and are not therefore considered to be an especially acquiescent group. Column 3 lists the average proportion of items in the total scale that were endorsed by this group ($\bar{p} = \bar{x}/n$). With the exception of the 566 "all true" scale (AT), which is not proposed as a meaningful acquiescence measure, none of the acquiescence measures depart from the expected average scale score value of $n/2$ ($\bar{p} = .50$).

TABLE 4

CHARACTERISTICS OF ACQUIESCENCE SCALES

Acquiescence scale	No. of items	$\bar{p}$	$\sigma_{\bar{p}}$	r_{KR-21}	$r_{SB-corr}$
Rb	27	.50	.12	.36	.56
Acq	24	.47	.13	.43	.67
Bn	32	.47	.12	.44	.60
At	18	.49	.16	.46	.75
B	63	.43	.10	.65	.65
AT	566	.41*	.06	—	—

* Different from $\bar{p} = .50$ at .001 level.

According to our previously developed rationale, a further requirement of a desirable acquiescent pool would be relatively low interitem correlations in this population of subjects. Since considerable care was taken in all of these item pools to insure that items would be of approximately equal difficulty level, Kuder-Richardson Formula 21 was employed to yield a lower-bound estimate of the internal consistency of the acquiescence scales. Column 5 presents the estimated internal consistency coefficients for acquiescence scales based on differing numbers of items. Unfortunately, we are unable to specify with any precision the degree of internal consistency an acquiescence scale should possess in this population. From a practical standpoint, it would seem that the .36 internal consistency coefficient of *Rb* represents the least amount of content variance that one might expect from an MMPI acquiescent item pool in a college group. The rather substantial internal consistency of Fricke's original *B* scale ($r = .65$) is almost identical with the .64 reported by Hanley (1961) for a college group and is taken as further confirmation of his suspicion that content variance is involved in this scale.

The last column of Table 4 gives the internal consistency coefficients corrected by the generalized Spearman-Brown formula to a common base of 63 items. Although this provides a common basis for comparison of the various methods of obtaining acquiescent items, it should be recalled that most of the methods were exhaustive. The 27 items that form the *Rb* scale, for example, represent the *only* items from the original 566-item pool that met the specified criteria for inclusion. The internal consistency of an *Rb* scale augmented by 36 items is therefore only of academic interest.

Acquiescence and Communality as Dimensions of the MMPI

There are several views on the dimensionality and relationship of agreement tendencies and communality that are not easily reconciled. Edwards (1957) originally took the position that the tendency to respond in terms of perceived social desirability value (communality) was a single dimension[4] and that agreement-disagreement tendencies were interpretable primarily in terms of this same concept. The present definition of acquiescence makes it logically independent of communality variance. Thus, with a pool of high controversiality items, the number of true responses (acquiescence) will be perfectly negatively correlated with the number of false responses (cautiousness). However, when items of extreme communality are involved, there is no necessary relationship between the proportion of "deviant true" responses and the proportion of "deviant false" since they are not required to sum to unity. Evidence that there is, in fact, no relationship would be Barnes (1956b) finding that Deviant True and Deviant False are uncorrelated in the total MMPI pool and the recent report (Jackson & Messick, 1961) that Deviant True and Deviant False items are negligibly correlated ($r = .13$) in the *F* scale.

Jackson and Messick (1958, 1961) have consistently called attention to the possibility of acquiescence being elicited differentially by item pools of differing levels of rated social desirability. Their most recent formula-

[4] Messick (1960) has recently cast serious doubt on the undimensional character of rated social desirability values.

tion (Jackson & Messick, 1961) of social desirability (communality) and acquiescence as *orthogonal* factors underlying MMPI scales is compatible with our assertion that there is no logical reason that these two measures should be related. Indeed, this factorial study provides the most direct evidence available for the independence of these two measures since their marker variable for acquiescence meets our restricted definition of this response style (Jackson & Messick, 1961).

Ideally, if the number of true and false deviant answers were neatly balanced in each MMPI scale, it would be possible to speak of relatively "pure" measures of noncommunality with respect to certain criterion group variables. That this state of affairs is unlikely, however, is indicated by the fact that 63% of the items in the MMPI reflect a lack of communality with the population at large when answered true while only 37% of the items indicate a divergence from population norms when answered false. On this basis alone, we would expect a strong deviant true bias to be present in the MMPI clinical scales and, indeed, Messick and Jackson (1961) have ably demonstrated that the strongest and most consistently isolated factor in MMPI factors studies is interpretable as a deviant true factor. Although Messick and Jackson (1961) speak of this factor as acquiescence, they are not using the term here in the present restricted sense of the word. We would consider this factor to be a confounding of acquiescent style and a method (test) that employs an unbalanced item pool.

Proponents of the Deviation Hypothesis (Barnes, 1956b; Berg, 1961) have postulated "response biases" that include both agreement tendencies and tendencies to answer deviantly (in the direction of noncommunality). The most convincing evidence for the existence of these two biases come from Barnes' (1956b) demonstration that the tendency to answer deviantly true on the MMPI is an impressive predictor of the "psychotic scales" while the tendency to answer deviantly false is substantially related to the MMPI "neurotic scales." Additional evidence of the importance of these postulated deviant response tendencies is found in Barnes' (1956a) later observation that Deviant True appears to be a "pure factor test" of Wheeler, Little, and Lehner's Factor I (1951) on the MMPI, while Deviant False is heavily loaded on Factor II of the same study.

Even if it is conceded that Deviant True and Deviant False appear to account for much of the variance in the MMPI, this does not commit one to the position that there are two independent "deviant response styles" in operation. These gross measures would seem to include all the sources of variation we have defined thus far, namely: variance in relative communality arising from the *strategy* of criterion groups, variance attributable to the postulated *stylistic* dimension of acquiescence-cautiousness, and variance attributable to the particular item pool represented by this *test* with its relative imbalance (63%) of low communality items keyed true.

In order to investigate the relationship of Deviant True and Deviant False to these other postulated sources of variation, it was thought necessary to replicate Barnes' (1956b) procedure in a slightly larger and more homogeneous population. In addition, we were interested in the relationships of the individual difference measures of social desirability (com-

munality) and acquiescence to these deviance dimensions as well as the relationships of several of the MMPI clinical and special scales which are relevant to the present argument.

The Deviation Hypothesis and the MMPI

With the foregoing in mind, a 356-item Deviant True key and a 210-item Deviant False key were constructed in a manner analogous to the scoring method used by Barnes (1956b) with individual record forms. The item scoring direction or direction of noncommunality for normative groups is given by the authors of the MMPI for all items (Hathaway & McKinley, 1951). Items were separated into Deviant True and Deviant False groups on this basis.

When scores from these two keys were correlated in our sample of 100 college subjects (50 men and 50 women), the correlation between them was found to be .001. This striking result fully supports Barnes' (1956b) contention that the two measures are independent of each other. In addition, the fact that Deviant True and Deviant False are completely uncorrelated leads to a satisfying clarity of conceptualization in viewing their relative contributions to possible stylistic dimensions and to the various MMPI scales. Figure 2 presents this conceptualization. The abscissa represents degrees of correlation of a given scale with the 210-item Deviant False key, ranging from −1.00 to

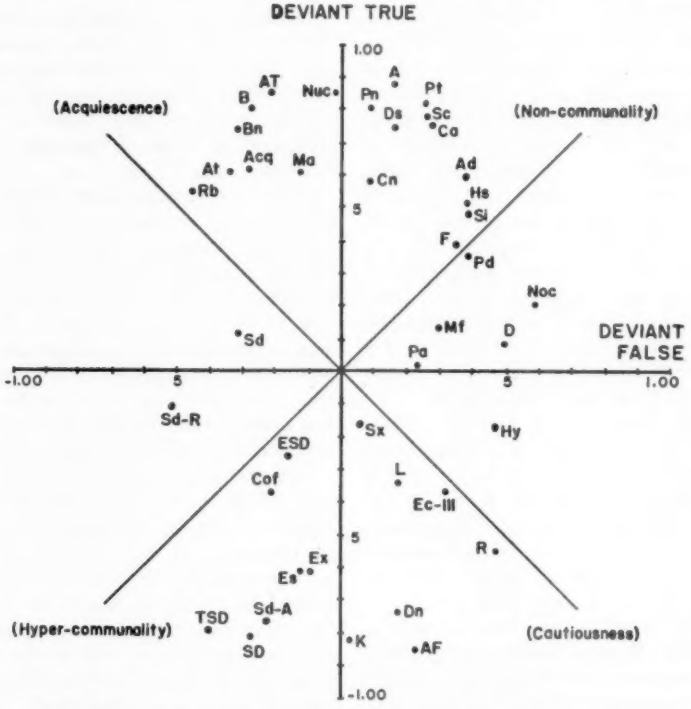


FIG. 2. Correlations of MMPI measures with Deviant True and Deviant False keys.

+1.00. The ordinate represents degrees of correlation of a given variable with the 356-item Deviant True key, from −1.00 to +1.00. The justification of this orthogonal plotting of the two response measures is the zero correlation that exists between them in this group of 100 college subjects.

A relatively pure measure of Deviant True is provided by Block's (1953) Neurotic Undercontrol scale (*Nuc*) which appears at the top of the y-axis. This measure of deviant undercontrol may be thought of as reflecting the "yea-saying" character style described by Couch and Keniston (1960). Welsh's (1956) Factor I or *A* scale is another strong candidate for inclusion as a measure of Deviant True as has been emphasized by Messick and Jackson (1961). It is of interest to note that while the *K* scale (at the negative end of the y-axis) is a strong measure of the extent to which a person does *not* respond in the Deviant True direction, it is unrelated to the tendency to give deviantly false responses. Pure measures of Deviant False are not as easily identified among the scales included in this analysis. However, it should be noted that Block's Neurotic Overcontrol scale (*Noc*), which is presumably orthogonal to his *Nuc* scale and which might be thought of as a "nay-saying" scale (Couch & Keniston, 1960), is highly related to the Deviant False dimension.

Perhaps the most interesting feature of this scatter plot is that the scales that fall within each of the four quadrants have a definite logical relation to one another. Quadrant I contains the MMPI clinical scales of the first factor variety, Quadrant II contains six different measures of response acquiescence, Quadrant III contains seven measures of social desirability, and Quadrant IV contains five scales that are presumably related to repressive overcontrol or denial.

The scales within a quadrant can be seen to vary in the extent to which they are related to Deviant True, Deviant False, or both response measures. The 45 degree lines extended out through the quadrants in a manner analogous to "fusion factors" (Kassebaum, Couch, & Slater, 1959) represent points of balance or symmetry between these two response measures.

That is, scales which fall along or near these lines may be considered to be *equally* influenced by both Deviant True and Deviant False keyings. Although the format of the MMPI and the purposes for which it was constructed make it unlikely that any relevant stylistic dimensions will be isolated that are unrelated to Deviant True and Deviant False, scales to which these measures contribute equally have some potential for measuring other stylistic dimensions. Consequently it is of interest to consider the scales that fall along these balance lines.

In Quadrant I, *Pd* and the *F* scale are quite close to the balance line. The *F* scale, which has been referred to as the "screwball scale" (Fricke, 1957) is a relatively powerful measure of noncommunality. It consists of items of highly asymmetrical endorsement frequencies which are keyed in the direction of deviance. In this particular college population, *Pd* is seen to be an even better-balanced measure of noncommunality. In Quadrant II, the distance of the various acquiescence scales from the balance line seems to reflect the extent to which care was taken in their construction to take out the effects of adjustment or communality variance. The most successful of the scales, in this respect, seems to be the *Rb* scale developed by the writer from a pool

of items of both high controversiality and neutral rated social desirability value. This scale is interpreted as reflecting the tendency to answer true to indifferent items or acquiescence as it has been used throughout this paper.

In Quadrant III, the social desirability scales tend to be heavily influenced by the tendency to respond in the opposite direction from Deviant True as represented by the *K* scale. Consequently, most of them are some distance from the line that would represent a balanced measure of hypercommunality. The *ESD* scale contains items of extreme social desirability value and it would probably be difficult to find individuals who did not score high on this scale. Cofer's scale (*Cof*) with its interesting item pool, already discussed, falls farther from the origin, but cannot be considered a completely "balanced" scale. The Quadrant IV balance line is well represented by Welsh's (1956) second factor "repression" scale, Scale *R*. Block's (1953) Ego Control scale (Ec-III), (which is presumably a fusion factor of the *Noc* and *Nuc* scales) also falls near this balance line. Scales which fall along this balance line are interpreted as having potential for measuring the tendency to deny relatively neutral or nondeviant statements in the MMPI ("cautiousness").

The point of view taken in this paper has been that Deviant True and Deviant False contribute to scale variance on the MMPI because they are the receptacles of variance due to strategy (differentiation of deviant criterion groups), method (the uniqueness of the MMPI item pool), and style (acquiescence-cautiousness). It was further assumed that scales which were found to be *equally* influenced by these two measures could be considered as having potential for the measurement of styles that are not peculiar to this instrument. By this reasoning, the *Pd* and *F* scales were interpreted as measures of the strategic dimension of communality, and the stylistic dimension of acquiescence-cautiousness was assumed to be represented by the *Rb* (response bias) scale at one end and the *R* (repression) scale at the other.

Extending this admittedly speculative reasoning even further, we can venture an interpretation of the possible psychological significance of Deviant True and Deviant False when considered in relation to the hypothetical dimensions of communality and acquiescence. The positive end of the Deviant True dimension which is represented by the *Nuc* would be considered to represent a fusion of acquiescence and noncommunality. The opposite end of Deviant True, represented by the *K* scale would be thought of as a fusion of hypercommunality tendencies with cautiousness. Although the Deviant False dimension is not purely represented by any scale, the *Noc*, *D*, *Hy*, and *Pa* scales could be thought of as fusions of cautiousness and noncommunality.

The above speculations, although plausible in the light of the known correlates of some of the scales discussed, have the status of intuitive hunches that must be subjected to more rigorous test. It is suggested that factor analysis is the method of choice and particularly of the type advocated by Jackson and Messick (1961). It is hoped that the distinctions made in this paper will guide the selection of appropriate "marker variables" in these analyses such that dimensions will be isolated which have generality beyond the MMPI as an instrument. The stylistic dimension of acquiescence and its com-

plement of cautiousness would seem to have some precedence as a theoretical construct. Likewise, the concept of communality, which stems from Berg's generally useful Deviation Hypothesis would seem to be worthy of further exploration. If the MMPI method imposes limitations on the development of stylistic measures that are independent of Deviant True and Deviant False, it would seem that efforts should be directed toward the development of balanced or equally influenced measures of other dimensions—and particularly that represented by the concept of hypercommunality.

SUMMARY

A distinction was drawn among three sources of variance that contribute to assessment with the MMPI. Strategic variance arises in the assessment of a subject's communality with respect to a normative group on a dimension defined by contrast with a criterion group. Method variance is due to the idiosyncratic nature of the total item pool in regard to the proportion of true and false keyings and the distribution of item popularity values. Stylistic variance includes dispositions to agree (acquiescence) or disagree (cautiousness) with neutral statements, independently of item content. With these distinctions as organizing concepts, research relating to social desirability, acquiescence, and deviant response biases in the MMPI was critically reviewed. Suggestions were made concerning the measurement of these variables within the MMPI so that factor analytic studies might be guided by marker variables of broadened theoretical significance.

REFERENCES

BARNES, E. H. Factors, response bias, and the MMPI. *J. consult. Psychol.*, 1956, **20**, 419–421. (a)

BARNES, E. H. Response bias in the MMPI. *J. consult. Psychol.*, 1956, **20**, 371–374. (b)

BERG, I. A. Response bias and personality: The deviation hypothesis. *J. Psychol.*, 1955, **40**, 61–72.

BERG, I. A. Deviant responses and deviant people: The formulation of the deviation hypothesis. *J. counsel. Psychol.*, 1957, **4**, 154–161.

BERG, I. A. The unimportance of test item content. In B. M. Bass and I. A. Berg (Eds.), *Objective approaches to personality assessment.* New York: Van Nostrand, 1959. Pp. 83–89.

BERG, I. A. Measuring deviant behavior by means of deviant response sets. In I. A. Berg and B. M. Bass (Eds.), *Conformity and deviation.* New York: Harper, 1961.

BLOCK, J. The development of an MMPI-based scale to measure ego control. Unpublished manuscript, Institute of Personality Assessment and Research, University of California, Berkeley, 1953.

COFER, C. N., CHANCE, JUNE, & JUDSON, A. J. A study of malingering on the MMPI. *J. Psychol.*, 1949, **27**, 491–499.

COHN, T. S. The relation of the *F* scale to a response set to answer positively. *Amer. Psychologist*, 1953, **8**, 335. (Abstract)

COUCH, A., & KENISTON, K. Yeasayers and naysayers: Agreeing response set as a personality variable. *J. abnorm. soc. Psychol.*, 1960, **60**, 151–174.

COWEN, E. L., & STILLER, A. The social desirability of trait descriptive terms: Order and context effects. *Canad. J. Psychol.* 1959, **13**, 193–199.

CRONBACH, L. J. Studies of acquiescence as a factor in the true-false test. *J. educ. Psychol.*, 1942, **33**, 401–415.

CRONBACH, L. J. Response sets and test validity. *Educ. psychol. Measmt.*, 1946, **6**, 475–494.

CRONBACH, L. J. Further evidence on response sets and test design. *Educ. psychol. Measmt.*, 1950, **10**, 3–31.

CROWNE, D. P., & MARLOWE, D. A new scale of social desirability independent of psychopathology. *J. consult. Psychol.*, 1960, **24**, 349–354.

DAHLSTROM, W. G., & WELSH, G. S. *An MMPI handbook: A guide to use in clinical practice and research.* Minneapolis: Univer. Minnesota Press, 1960.

EDWARDS, A. L. The relationship between the judged desirability of a trait and the

probability that the trait will be endorsed. *J. appl. Psychol.*, 1953, **37**, 90–93.

EDWARDS, A. L. *The social desirability variable in personality assessment and research.* New York: Dryden, 1957.

FISKE, D. W., & RICE, L. Intra-individual response variability. *Psychol. Bull.*, 1955, **52**, 217–243.

FRICKE, B. G. A response bias (*B*) scale for the MMPI. *J. counsel. Psychol.*, 1957, **4**, 149–153.

FULKERSON, S. C. An acquiescence key for the MMPI. *USAF Sch. Aviat. Med. Rep.*, 1958, No. 58–71.

GAGE, N. L., LEAVITT, G. S., & STONE, G. C. The psychological meaning of acquiescence set for authoritarianism. *J. abnorm. soc. Psychol.*, 1957, **55**, 98–103.

GUILFORD, J. P. *Psychometric methods.* (2nd ed.) New York: McGraw-Hill, 1954.

HANLEY, C. Social desirability and responses to items from three MMPI scales: *D*, *Sc*, *and K. J. appl. Psychol.*, 1956, **40**, 324–328.

HANLEY, C. Deriving a measure of test-taking defensiveness. *J. consult. Psychol.*, 1957, **21**, 391–397.

HANLEY, C. Social desirability and response bias in the MMPI. *J. consult. Psychol.*, 1961, **25**, 13–20.

HATHAWAY, S. R., & MCKINLEY, J. C. *The Minnesota Multiphasic Personality Inventory manual.* (Rev. ed.) New York: Psychological Corporation, 1951.

JACKSON, D. N., & MESSICK, S. Content and style in personality assessment. *Psychol. Bull.*, 1958, **55**, 243–252.

JACKSON, D. N., & MESSICK, S. Acquiescence and desirability as response determinants on the MMPI. *Educ. psychol. Measmt.*, 1961, **21**, 771–790.

KASSEBAUM, G. G., COUCH, A. S., & SLATER, P. E. The factorial dimension of the MMPI. *J. consult. Psychol.*, 1959, **23**, 226–236.

MEEHL, P. E. The dynamics of "structured" personality tests. *J. clin. Psychol.*, 1945, **1**, 296–303.

MEEHL, P. E., & HATHWAY, S. R. The *K* factor as a suppressor variable in the MMPI. *J. appl. Psychol.*, 1946, **30**, 525–564.

MESSICK, S. Dimensions of social desirability. *J. consult. Psychol.*, 1960, **24**, 279–287.

MESSICK, S., & JACKSON, D. N. Acquiescence and the factorial interpretation of the MMPI. *Psychol. Bull.*, 1961, **58**, 299–304.

SECHREST, L., & JACKSON, D. N. Deviant response tendencies: Their measurement and interpretation. Paper read at American Psychological Association, Chicago, September 1960.

WELSH, G. S. Factor dimensions *A* and *R*. In G. S. Welsh and W. G. Dahlstrom (Eds.), *Basic readings on the MMPI in psychology and medicine.* Minneapolis: Univer. Minnesota Press, 1956.

WHEELER, W. M., LITTLE, K. B., & LEHNER, G. F. J. The internal structure of the MMPI. *J. consult. Psychol.*, 1951, **15**, 134–141.

WIGGINS, J. S. Interrelationships among MMPI measures of dissimulation under standard and social desirability instructions. *J. consult. Psychol.*, 1959, **23**, 419–427.

WIGGINS, J. S. Definitions of social desirability and acquiescence in personality inventories. In S. Messick and J. Ross (Eds.), *Measurement in personality and cognition.* New York: Wiley, 1961.

WIGGINS, J. S., & RUMRILL, C. Social desirability in the MMPI and Welsh's factor scales *A* and *R*. *J. consult. Psychol.*, 1959, **23**, 100–106.

(Received January 16, 1961)

Psychological Bulletin
1962, Vol. 59, No. 3, 243–247

COMPARATIVE PSYCHOLOGICAL STUDIES OF NEGROES AND WHITES IN THE UNITED STATES:

A CLARIFICATION

BENJAMIN PASAMANICK
Columbus Psychiatric Institute and Hospital, Ohio

Dreger and Miller (1960) in their recent review of Negro-white differences dismiss our early work on Negro child development (Pasamanick, 1946) with the blanket indictment that "Inadequacies of sampling, as well as the more general difficulty of estimating skin color subjectively, tend to invalidate Pasamanick's conclusions."

To clarify for ourselves and the readers of the review the basis for the general rejection of our study we entered into correspondence with the senior author (Dreger) seeking an elaboration and specification for his criticism. The result of this correspondence was the expansion of the condemnatory statement into six specific objections. Two concerned sampling methods, one dealt with the infant examination procedures, another with the possibility of inherent bias in the examiner, one with the method of dichotomizing the Negro infants by skin color as light or dark, and finally one which questioned the reliability and validity of the Gesell developmental methods used to obtain the DQ for our subjects.

The first and most important objection involves the contention by Dreger that the Negro infants in the study were born to parents who were too well educated to be representative of the New Haven Negro population. Dreger calls attention to the fact that in 1946 the median years of schooling for the entire United States population, 25 years of age and older, was 8.6 years and that by 1950 it had only risen to 9.3 years. For the United States Negro population the median years of schooling was even lower. For example, in 1940 the median years of schooling for Negroes, age 25 or over, was only 5.7; even in Connecticut the Negro median was only 7.6. (Census data taken from the United States Bureau of the Census, 1952a, 1952b). Hence, since the parents of these infants had attained over 10 years of schooling on the average Dreger believed they could hardly represent the Negro population in Connecticut or elsewhere.

This objection can be answered easily. The error lies in lumping together the educational attainments of all Negores 25 years of age and over, thereby including the older and far less well educated individuals. If one restricts the analysis to New Haven and to the age category of the parents of the Negro infants, i.e., the child-bearing years of 20–40, it becomes readily apparent that they do indeed typify the New Haven Negro population within this age range. Further, it is our contention that the educational attainment of parents is largely irrelevant as far as the DQ of infants is concerned. It was indicated in the study criticized by Dreger that children whose parents had an educational level little more than grammar school did not differ in mean scores from the children of parents who had approximately a high school education. We confirmed these findings in the 1,000 subjects involved in our Baltimore study of

child development and again found that education of parents as well as socioeconomic status apparently does not play an important role in either whites or Negroes until the children are 3 years of age or older (Knobloch & Pasamanick, 1960).

The second Dreger objection is of the same order. He suggested that the white infants in our study were also unrepresentative of white infants generally. Specifically he cited our inclusion of 61 white illegitimate infants (an adoption group), nursery applicants whose parents averaged over 17 years of schooling, and a third group comprising residents of an orphan asylum.

In response to this objection it should be noted that the study population of white infants was not selected as being representative of New Haven white infants. *The major comparison in our paper involved a contrast of the aforementioned representative Negro infants with the normative group used by Gesell for establishing his developmental norms.*

The white infants in our investigation were selected for reasons quite apart from this comparison which were made explicit in the paper. For example, because of a Jamaica study on Negro children from a crèche for whom quite low scores were obtained we included the institutionalized group of white infants. Our adoption group was included in order to compare the illegitimate children in and out of an institution. Infants of white upper and middle class parents were studied because of our feeling that, regardless of the class standing of parents, white and Negro infants do not differ appreciably in behavioral development during early life. These three groups, the upper class white, the adoption, and the illegitimate, were specifically *excluded* as controls. They represented substudy populations and served only for comparative purposes. They were, in short, three specially selected criterion groups.

Dreger's third criticism is based largely on a misconception of our infant examination procedure. The objection, simply stated, is that Negro mothers most often were present and white mothers most often absent during the examination of the infants. Since the presence of the mother may be presumed to influence positively the response of the infants he considered that the Negro infants were unduly favored.

The actual procedure which was stipulated and enforced, and is indeed standard Gesell method, was that the mothers or mother substitutes, e.g., the nurse who cared for the institutionalized infants, be present during the course of the examination. Such was the case for all infants, white and nonwhite.

The fourth contention of Dreger is that examiner bias could have been a potent factor in uprating the scores of the Negro infants. The presumption is that such bias would of necessity have favored the nonwhite infants. The further presumption is that such bias could not be adequately controlled. We, of course, are compelled to agree that such bias is a possibility and further that there is no known method of assessing or controlling its influence. On the other hand there is no reason to suspect that its presence in either direction was greater in this than in other investigations of the same type. That it might not even be as great was indicated in an implicit test of bias in the Baltimore replication. In that study the examiner, who did not know the status of the subjects, found lower quotients and more

damage in the premature than in the control infants, both Negro *and* white (Knobloch & Pasamanick, 1958).

In his fifth specific criticism Dreger pointed out that dichotomizing the skin color of Negro infants into light and dark is at best subjective and probably unreliable as well. He argued that such color distinctions are not our problem alone but concluded with the admonition that since we dealt with subjects "as close to being only hereditarily influenced as it is possible to get" the problem is presumably more critical in our study than might be the case in other studies later in life.

Again we agree wholly with the argument that skin color determinations are both subjective and often unreliable. For precisely these reasons great pains were taken to examine and standardize our procedures for skin color determination. At first a color wheel was utilized. Later two judges were found to agree with one another in the determination of light and dark skin with almost perfect reliability. Despite this reliability we are perfectly willing to grant that skin color is not very good evidence of the degree of mixture of white and nonwhite strains. Since others found significant differences when making such a comparison we felt it warranted and necessary to replicate their work and utilized the best techniques currently available in doing so. Incidentally, we found no significant differences.

The sixth criticism of our paper by Dreger centers upon his dissatisfaction with the Gesell schedules, presumably on the score of reliability and validity.

In respect to this point which would require an undue amount of space to discuss we would prefer to refer the reader to papers which comment on the adequacy of the Gesell schedules and indicate that in trained hands correlations of .5 to .75 can be secured with the infant scores 2–7 years later. In addition, interjudge reliability with correlations centering about .9 are obtained (Knobloch & Pasamanick, 1960, 1961).

In reply to the counter discussion given above Dreger wrote the following: (It is given verbatim since we are unable fully to comprehend the criticism. Nevertheless our attempt at clarification, witl in the limits of our comprehension of the points made, follows his comments.)

> Your point is well taken that your Negro infants were representative of New Haven infants of their race. I cannot see, however, that they are representative of Negro infants throughout the country; if they were not intended to be so, considerable misunderstanding has existed in my own mind—and others', too, I believe. It also seems that parental educational level does have something to do with representativeness of infancy. One of the contentions of hereditarians is that, of course, there are wide individual differences among Negroes; if then comparisons of whites from different levels are made with Negroes from the upper levels, an unfair conclusion is drawn. As you can see, a modification of the selective migration hypothesis is involved here. However, if the major comparison you made was between what amounts to upper level educational Negroes (parent level) and Gesell's normative group, I should judge no harm would result. For even though you point out that his norms are more representative than your white samples, they seem to be almost certainly from upper classes. At least, that is the general criticism I have heard of Gesell's norms across the years. One further point: I am not clear as to the irrelevancy of socio-economic status in your original article. Your further work in Baltimore appears to substantiate your claim that these variables do not exert their effect until later. But the fact that they were not controlled for in the earlier work might make it necessary to conclude from that study alone that the lack of differences between the two groups could be the result of socio-economic factors.

We are somewhat at a loss to

understand the criticism that while our Negro infants were representative of New Haven infants of their race they are not representative of Negro infants throughout the country. Since, as far as we are aware, no representative group of Negro children in the United States has ever been compared with any representative group of white children for the country as a whole, the *Psychological Bulletin* review as a whole would appear to lack relevance. It would merely have sufficed to state that all the groups studied were unrepresentative and that only local conclusions could be drawn but no generalized statements were warranted.

Of course, the New Haven Negroes studied could not be claimed to be representative of all Negroes in the United States, although I dare say they are probably not too unrepresentative. The important point was the similarity to whites during infancy of a group who later in life would almost certainly display intelligence tests scores significantly lower than the white scores of comparably aged whites. The similarity existed despite the fact that the children came from homes that were lower socioeconomically and educationally.

Precisely the same findings were described in our Baltimore investigation where hundreds of Negro and white children were involved. In these studies special efforts were made to see to it that both racial groups were representative of their respective populations. By almost all demographic variables Baltimore Negroes are probably as representative of American Negroes as can be found in any disadvantaged group. Selective migration would not be expected to play any important role in this city south of the Mason-Dixon line. In our study the *N* was large enough so that internal comparisons could be tested adequately. As stated previously, neither race, parental education, or socioeconomic status during infancy, was associated with any differences in infant performance.

A point is made that the Gesell norms were derived from upper classes and that this is the general criticism heard by Dreger across the years. Hearsay is, of course, not the best evidence. Reference to Gesell's (Gesell & Thompson, 1938) description of the first normative sample indicates that it was drawn from a fairly homogeneous lower middle class group. The norms were changed slightly in "Developmental Diagnosis" (Gesell & Amatruda, 1941) on the basis of a heterogeneous group never described in the literature but not very different from the original sample. However, I must point out that had the norms been derived from an upper socioeconomic group (one such group is included in the paper under criticism) one would presumably expect that these norms would be higher than those derived from a lower class group. Accordingly then, following Dreger's logic even greater differences might have been expected in comparisons with the lowest American socioeconomic group, i.e., Negroes. However, the fact that no differences at all were found should suffice to indicate that Dreger's statement "that the lack of differences between the groups could be the result of socio-economic factors" is quite irrelevant.

We are grateful for the opportunity afforded us to reply to the crucial questions raised by the Dreger and Miller critique of our work and to Dreger for his generous permission to use our correspondence. It might appear that the authors were a bit premature in disposing of our work.

REFERENCES

DREGER, R. M., & MILLER, K. S. Comparative psychological studies of Negroes and whites in the United States. *Psychol. Bull.*, 1960, **57**, 361–402.

GESELL, A., & AMATUDA, C. S. *Developmental diagnosis.* New York: Hoeber, 1941.

GESELL, A., & THOMPSON, H. *The psychology of early growth.* New York: Macmillan, 1938.

KNOBLOCH, H., & PASAMANICK, B. The relationship of race and socio-economic status to development of motor behavior patterns in infancy. In B. Pasamanick and P. Knapp (Eds.), *Social aspects of psychiatry.* (Regional Research Reports, No. 10) Washington, D. C.: American Psychiatric Association, 1958. Pp. 123–133.

KNOBLOCH, H., & PASAMANICK, B. Environmental factors affecting human development, before and after birth. *Pediatrics*, 1960, **26**, 210–218.

KNOBLOCH, H., & PASAMANICK, B. The consistency and predictability of the Gesell forty-week developmental schedules. In C. Shagass and B. Pasamanick (Eds.), *Child development and child psychiatry.* (Regional Research Reports, No. 13) Washington, D. C.: American Psychiatric Association, 1961.

PASAMANICK, B. A comparative study of the behavioral development of Negro infants. *J. genet. Psychol.*, 1946, **69**, 3–34.

UNITED STATES BUREAU OF THE CENSUS. *Census of population: 1950.* Vol. 2. *Characteristics of population.* Part 7 (Connecticut). Washington, D. C.: USBC, 1952. (a)

UNITED STATES BUREAU OF THE CENSUS. *Census tract statistics: 1950 population census report.* Vol. 3. Ch. 35. Washington, D. C.: USBC, 1952. (b)

(Received January 18, 1961)

Psychological Bulletin
1962, Vol. 59, No. 3, 248–256

METHODOLOGICAL ISSUES IN *P* TECHNIQUE[1]

WAYNE H. HOLTZMAN
University of Texas

During the past 15 years, increasing attention has been given in psychology to the intensive study of the single individual with repeated measures gathered over a long period of time. For the most part, intra-individual differences due to repeated testing of the same person have been viewed as error variance to be ignored or canceled out by some kind of averaging process. Most of psychology has been primarily concerned with the development of nomothetic laws governing behavior in general in which the uniqueness of the individual is completely lost. It is only recently with the advent of new conceptual tools and high speed computers that the statistical study of the single case has been seriously attempted.

Searching for a systematic approach to the study of individual differences, Cattell (1946) developed a scheme that he has called the covariation chart, consisting of major reference axes upon which a taxonomy of experimental designs can be constructed. Originally consisting of three primary axes, the covariation chart could be outlined as a box-like structure with scores across different subjects, different tests, and different occasions representing the three dimensions. *P* technique was coined by Cattell to denote the special case where the intercorrelations among *K* scores across *T* occasions are factor analyzed for a single individual.

[1] Paper presented at Symposium on the Development of Multivariate Experiment in Psychology, University of Illinois, November 14–16, 1960.

The Model of Multiple Time Series

Stripped of its psychological connotations, the basic model for *P* technique has been known by the statistician as multiple time series. A series of observations taken in a definite time order usually consists of measures which are not independent of one another and which consequently do not satisfy the assumption of independence made in most statistical analyses. The interval between successive observations is arbitrary, usually a day, a month, or a year, depending upon the preference of the investigator or the convention of society. Everyone is familiar with such time series as the cost-of-living index, annual population estimates, and monthly rainfall. Because time is the primary dimension for the ordering of variables and the definition of the unit of measurement, the variety and number of plausible variables are almost infinite. Of course in any one study, it is only possible to include a small number of time series, hopefully those observations most likely to cast some light on the dynamic nature of the process.

Traditionally, the greatest amount of attention has been given to the analysis of the internal structure of one or more time series in order to understand the factors which influence a particular variable such as the size of population. The forecasting of future trends from naturalistic observations has grown into a complex empirical science involving the search for statistical regularities in time series. More recently, a some-

what different emphasis has been developed by a growing number of statisticians, namely, the multivariate analysis of correlations between time series. It is this latter focus which is of central interest to the psychologist employing *P* technique.

In its simplest form, *P* technique involves no more methodological difficulties than any other factor analysis. The score distributions are assumed to be reasonably normal with complete independence of observation from one point in time to the next. The fact that variance is measured across time rather than across individuals makes no difference if indeed the successive observations in a given time series are independent. One has only to repeat the observations over a sufficiently large number of occasions to gain stability in the intercorrelations and resulting factor analysis. It would make the life of a psychologist or statistician much easier if this simple model for *P* technique actually worked. Unfortunately, all too often the observations made on a single individual through time are highly correlated rather than independent. Let us take a closer look at the nature of such serial dependence and what can be done to allow for it in the analysis.

The Problem of Serial Correlation

Perhaps the most serious problem from the psychologist's point of view is the effect of repeated measurement upon that which is being measured. One is restricted in *P* technique to those variables which can be measured again and again on the same person without serious loss of validity. This very point means that there must be wholesale elimination of many important psychological variables simply because there is no way to measure them repeatedly. It may be possible to develop abbreviated alternate forms of some tests in sufficient quantity to make feasible their use for repeated measurement of the same individual. Moran (1959) has done an admirable job of developing a special battery of repetitive psychological measures for the study of mental ability. But in general, the range of mental tests which can be adapted to repetitive use is quite limited.

Given that a variable can be measured repeatedly, there still remains the problem of sequential effects through time. For any variable which involves strong practice, fatigue, or adaptation effects through time, the degree of serial correlation will be high and must be taken into account in any analysis of the variable. The amount of such correlation between serial observations may vary considerably from one phase in the time series to another, as well as from one variable to the next, complicating the analysis still further.

The most direct method for determining the amount of sequential dependency of observations in a time series is to compute the serial correlation coefficients for the series. For the first serial correlation coefficient, usually symbolized as r_1, the first observation is paired with the second, the second with the third, and so on until the last observation in the series has been reached. Similarly, the second, third, and nth coefficients can be computed by appropriate lag between the two observations constituting the pairs being correlated. The entire array of serial correlation coefficients make up a correlogram when arranged in graphic form with the magnitude of the correlation expressed as a function of the lag. The significance of any serial correlation

may be tested approximately by using any table for ordinary correlation coefficients.

Correlation coefficients may be computed between time series in the usual manner, regardless of the degree of serial correlation present in the two series. Interpretation of such cross-correlations, however, is another matter. In the first place, the stability of the cross-correlation is directly affected by the degree of serial correlation present in the two time series. The degrees of freedom for testing the significance of the cross-correlation are reduced due to the sequential dependency of observations.

For example, suppose two variables, x and y, are measured daily for 100 days in an experiment using P technique. The cross-correlation between x and y is .30; the first four serial coefficients for x are .70, .50, .40, and .20; and the same coefficients for y are .50, .40, .30, and .10. Following Quenouille's suggestion (1952, p. 170), Bartlett's formula may be used for estimating the number of observations equivalent to one independent observation. For the present example, the number of paired observations needed to equal one degree of freedom for purposes of statistical inference is *at least* 2.38. If serial correlation beyond the fourth coefficient had been included in the equation,

$$1 + 2r_{x1}r_{y1} + 2r_{x2}r_{y2} + 2r_{x3y3} + 2r_{x4y4} + \cdots = 2.38 + \cdots$$

the estimate would have been still higher. In other words, there are fewer than 42 degrees of freedom, instead of 98, for testing the significance of the obtained correlation, too few for a correlation of .30 to prove significant beyond the .05 level. If the serial correlations had not dropped to zero after the fourth lag, the reduction in degrees of freedom would have been even more severe.

A second major difficulty obscuring the meaning of cross-correlations when there is serial correlation in the two time series being correlated, arises from the fact that significant lag correlation may also exist between Variable x at one point in time and Variable y at a later or earlier point in time. In the above example, the cross-correlation with zero lag is .30. But what about the cross-correlation when the observations on x at time t are paired with the observations on y at time $t-1$, $t-2$, or $t-s$?

Quenouille (1952, p. 170) has suggested the use of partial correlation coefficients to gain insight into the relationship between x and y. A given time series can be thought of as containing two parts: a systematic part which is predictable at any point in time from the preceding observations, and a random or residual part after removal of the serial correlation. The cross-correlation between simultaneous random parts in the two series can be obtained in a straightforward manner by partialling out the systematic parts of both variables. If one is interested in determining whether the nonsimultaneous random parts of two series are correlated, lagged partial correlation coefficients have to be calculated.

More recently, Quenouille (1957) has expanded upon this method for the analysis of multiple time series. His technique first involves the computation of complete intercorrelation matrices among the K time series, including those obtained when each series is lagged in time behind the others. The second step is to determine whether a Markoff scheme is operating, and if so, what order scheme is necessary to explain the systematic components in the K time

series. The next step is to examine the latent roots of the quotients $R_s R_{s-1}^{-1}$, where R_s is the intercorrelation matrix of cross-correlations lagged s intervals in time, in order to determine the specific nature of any trend components and relationships independent of trend. In addition to requiring excessively laborious computations, Quenouille's model for the analysis of multiple time series is still in a strictly experimental stage of development. Nevertheless, it serves to dramatize the difficulties one may encounter in dealing with multiple time series of the type often included in P technique experiments.[2]

The Use of Factor Analysis in P Technique

The extent to which the presence of serial correlation distorts the factor analysis of zero-lagged cross-correlations in multiple time series—the usual intercorrelation matrix for P technique—is not known. Most investigators have been content to ignore the problem of serial correlation. Others have dealt with it in only a very crude manner by adding synthetic time variables to the correlation matrix and eliminating trends by the rotation of factors on which the time variables are loaded. For example, Mefferd, Moran, and Kimble (1958) included three synthetic time variables in an analysis of 44 time series, each containing 144 observations. The first was intended to account for any linear trend and consisted merely of the day number ranging from 1 to 144. The second was designed to pick up any rapid cycling effect over a 48-hour period and consisted of the digits 1, 2, 1, 2 . . . beginning with the first set of observations and running through to the last. The third synthetic variable was also an oscillating one, but with a slower period consisting of the digits 1, 1, 2, 2 . . . running through the series. One last suggestion that has not yet been tried out successfully was put forth by Demarin (1958) who proposed the use of Quenouille's partial correlation method to correct cross-correlations for serial correlation prior to carrying out the factor analysis.

In a critical review of P technique, Anderson (1958) questioned seriously the advisability of any of these methods for dealing with systematic effects, particularly the introduction of errorless, synthetic time variables which are then treated by factor analysis in the same manner as the other measures. Although Anderson has no general method to offer in lieu of factor analysis for dealing with large systems of P technique data, he does recommend that such trend components be eliminated directly by regression techniques prior to the factor analysis. But as Cattell (1961) has pointed out, in the usual study where there are many variables, factors that are highly correlated with time will separate from one another

[2] Another slightly different approach to multiple time series has been taken by G. P. Williams and E. K. Harris who have examined the effect of serial correlation on principal components. "The system of auto- and cross-correlations among n variables may be expressed in the form of a symmetric matrix, consisting of circulant submatrices, each of which contains the lagged correlations between a given pair of variables. The eigenvalues, eigenvectors, and inverses of such submatrices are known. An analytical solution for the eigenvalues of this matrix has not yet been obtained. However, to illustrate the problem, four meteorological variables (temperature, relative humidity, barometric pressure, and wind speed) have been used. Each series consists of daily readings over a five-year period." (From abstract of paper presented at the Annual Meeting of the American Statistical Association, Chicago, Illinois, December 30, 1958.)

without serious distortion if rotation to simple structure is achieved. Unfortunately, the necessary theoretical analysis and empirical determination of the conditions under which trend components due to serial correlation distort the factor analysis or render it invalid have yet to be done.

A second major weakness in most factor analytic studies employing *P* technique is failure to examine anything but the zero lag cross-correlations between the time series. Most investigations involving repeated measurement on a single individual include several systems or levels of behavior in the same analysis. These systems may range all the way from biochemical assay of metabolic products in urine to behavior ratings in social situations. In the comprehensive study by Mefferd, Moran, and Kimble (1958), for example, six different systems of time series—chemical, physiological, psychological, psychiatric, environmental, and synthetic—were combined in one analysis. It is highly reasonable to expect some time lag in the relationships of variables in one system relative to those in another. Biochemical changes may precede or follow changes in the physiological, mental, or behavioral domains. Indeed, as Cattell, Mefferd, and others have pointed out, it is more plausible to hypothesize some time lag between certain variables than it is to deal only with the correlations across simultaneous observations.

Cattell (1958) has suggested that at least the first two or three lag-correlations be computed systematically for all cross-correlations. The largest correlation coefficient (what Cattell calls the "maximized lead-and-lag correlation") should then be inserted in the correlation matrix as the best estimate of r_{xy} for subsequent factor analysis. Anderson (1958) objects to this procedure on two grounds: only the strongest relationship goes into the factor analysis, completely glossing over the other relationships; and the original factor analytic model is no longer applicable because the degree of lag varies from one cross-correlation to the next.

A third criticism of Cattell's procedure can be made which recognizes a difficulty which is rather general in *P* technique. Rarely is it possible to obtain a very large number of observations through time in psychological studies of a single person. The presence of serial correlation reduces still further the stability of obtained correlations. When one considers the large number of cross-correlations possible, the likelihood of severe distortion due to exploitation of chance is very high. In the data obtained by Mefferd and his colleagues, for example, there are 6,622 cross-correlations if one allows a lag varying only from 0 to 3. The number of additional cross-correlations increases at the rate of 1,892 per lag. Since the degrees of freedom for evaluating the stability of any one of these coefficients will be somewhat less than 142, depending upon the degree of serial correlation present in the two time series, the dangers of Cattell's post hoc method of selecting the maximum cross-correlations should be apparent. Of course, where one has a good a priori reason from previous experimental studies to assume that a given lag between x and y is appropriate, one has a basis for selecting certain lagged correlations for special analysis.

The presence of clearly distinctive systems of variables or levels of human functioning in the same set of repeated measures over time suggests that there may be better methods of multivariate analysis than straight-

forward factor analysis in which all *K* variables are treated in an identical manner. The correlations between systems of variables are far more interesting than the correlations within systems. In those cases where such distinctive systems exist a priori, as in the time series collected by Mefferd, Moran, and Kimble (1958), Tucker's (1958) interbattery method of factor analysis appears to be more appropriate than traditional procedures. Of course, if one really makes a thorough analysis of multiple time series along the lines suggested by Quenouille (1957), only a little extra work is required to calculate the canonical correlations and variates that emerge from the relationships between the systems. In this way it would be possible to maximize most efficiently the nonvanishing correlations between canonical variates within such systems as the biochemical measures from urinary metabolites, the psychophysiological response variables, and the scores on tests of mental ability.

Special Problems Encountered in Psychophysiological Variables

Cattell has made a useful distinction between trait and state with respect to the appropriate multivariate design to employ a given experiment. The distinction can best be illustrated by reference to the concept of anxiety. As Cattell states (1961):

> Popular speech recognizes that one can have an anxious person—a person who all his life is characterologically operating at a higher anxiety level—and a *typically* non-anxious person who is *temporarily* in a highly anxious *state*. Is there a continuum in nature and form of expression between permanent (trait) and momentary (state) anxiety? That is to say, is characterological or trait anxiety just temporary state anxiety held permanently high or are trait and state different things with different patterns (Ch. 9, p. 149)?[3]

Recognizing the possibility that a given measurement applied once to a sample of individuals may not involve the same functional unities as the identical measurement applied to one individual over a number of occasions, Cattell has recommended that systematic programs of research be undertaken in which a number of marker variables from *R* technique experiments across individuals is included in *P* technique experiments across occasions. Presumably, then, by careful matching, factors emerging in studies of the single individual can be more clearly identified by reference to the marker variables from earlier nomothetic studies. While I do not share Cattell's optimism that there exists only a limited set of 10–20 response dimensions built into the repertoire of the human being and that these will emerge in both *R* and *P* technique experiments, the basic idea of common marker variables is a good one.

In dealing with psychophysiological measures, special problems arise which make it very difficult to develop appropriate marker variables for inclusion in both types of investigations. Indeed, it is difficult enough to decide which of many possible dependent variables from a given measuring instrument is the most appropriate to use, whether one is studying many individuals or only one. Too frequently the choice of a particular dependent variable is dictated by the peculiar characteristics of the apparatus employed rather than by any firm confidence in the validity of the measure with respect to some concept such as anxiety or autonomic reactivity.

The work by Holtzman and Bitter-

[3] Raymond B. Cattell and Ivan H. Scheier, *The Meaning and Measurement of Neuroticism and Anxiety*. Copyright 1961, The Ronald Press Company.

man (1952, 1956), Bitterman, Krauskopf, and Holtzman (1954), and Bitterman, Holtzman, and Barry (1955) on the development of measures from the conditioning and extinction of the galvanic skin response is a good example of the problems one typically encounters in deriving a set of meaningful psychophysiological variables worthy of systematic comparison with other data in a factorial study. Problems of apparatus design, choice of experimental procedures, control of physical environmental and apparatus variables, units of measurement, scaling, correction for baseline performance, elimination of both experimental and statistical artifacts—all plague the investigator from start to finish if he takes seriously the mandate to employ reliable, valid procedures that are multivariate in nature.

Unlike such molar variables as mental test scores or behavioral ratings, psychophysiological variables are often continuous in nature. The usual procedure is to employ some kind of continuous recording apparatus such as an electrocardiograph, electroencephalograph, or dermohmeter for measuring changes in skin resistance. A prearranged set of stimuli may follow a period of adaptation and, in turn, may be followed by a period of recovery to a resting state. Simultaneous recording of several physiological measures is often undertaken. The entire time of measurement may take only a few minutes and may be repeated on a number of different occasions as in *P* technique studies. Viewed conceptually, such a program of data collection can be thought of as successive sets of multiple time series on the single individual, one set for each separate occasion. Scores generated from time series analysis in the microcosm of the single occasion, in turn become the bits of information constituting the multiple time series in the macrocosm of the single individual across many occasions, out of which, in turn, is generated the supermacrocosm across many individuals. Is it any wonder that one may well get discouraged in attempts to map out a grand scheme for the multivariate description and explanation of human behavior!

The selection and quantification of variables in the microcosm of time series collected on a single occasion has been studied recently by communication engineers, biophysicists, and statisticians, employing the methods of spectral analysis, autocorrelation, and the theory of random processes. An excellent overview of selected work on the problems of data reduction when dealing with neuroelectric phenomena is given by Rosenblith and his associates in the Communications Biophysics Group at the Massachusetts Institute of Technology (1959). In discussing the use of correlation techniques for the reduction of EEG data, Molnar, Weiss, and Geisler (1959) adopt a random process model and analyze both autocorrelation and cross-correlation functions over short periods of multichannel recording repeated periodically over many days on a single individual, essentially the same design as *P* technique. For the immediate future, however, it is more likely that the investigator conducting *P* technique experiments will have to fall back on his intuition or make arbitrary choices with respect to the appropriate units of measurement to employ when dealing with psychophysiological variables through time.

Concluding Remarks

Nothing has been said about the methodological issues that are common to all forms of factor analysis,

regardless of whether the correlation matrix is obtained from *R* technique across subjects or from *P* technique across occasions. Enough is being said about these more general problems by other participants in this symposium. It is sufficient here to point out that these other issues do exist and must also be taken into account in *P* technique experiments. No attempt has been made to summarize, review, or criticize specific studies involving *P* technique, although several dozen such studies have been reported in the past decade. Rather, I have chosen to stick closely to those methodological difficulties peculiar to *P* technique that are as yet largely unsolved. Until much further work of a fundamental nature on the theory of analysis of multiple time series, it is unlikely that major improvements in *P* technique will be made. It will take an unusual combination of mathematical-statistical sophistication and first-hand familiarity with the subtleties of empirical data gathered by repeated measurement of the single individual to solve the frustrating problems introduced by the serial correlation present in most time series.

In the meantime, what is the investigator to do? While recognizing many of the issues outlined above, Cattell, Mefferd, and others insist that factor analysis is still a reasonably good tool for the reduction and interpretation of data gathered through repeated measurement of a single individual. The model of *P* technique is appealing in many respects. Most psychologists would admit that even a crude statistical method for studying the uniqueness of the single individual is better than none at all.

The situation may not be as gloomy as implied in the above discussion. There are indeed some psychological, physiological, biochemical, and behavioral variables that have little or no serial correlation when measured repeatedly through time. Certainly the currently advocated procedures of factor analysis are reasonably valid in such instances. Given extensive computer resources, the mathematical sophistication to follow some of the latest methods for analysis of multiple time series, such as those advocated by Quenouille, and a healthy skepticism concerning the stability of any results on a single case unless replicated again and again, there is no reason why one shouldn't move ahead immediately with such investigations, even where a high degree of serial correlation is present.

REFERENCES

ANDERSON, T. W. Comments on factor analysis of multiple time series. Paper presented at Symposium on multivariate analysis of repeated measurements on the same individual, American Psychological Association, Washington, D. C., September 1958.

BITTERMAN, M. E., KRAUSKOPF, J., & HOLTZMAN, W. H. The galvanic skin response following artificial reduction of the basal resistance. *J. comp. physiol. Psychol.*, 1954, **47**, 230–234.

BITTERMAN, M. E., HOLTZMAN, W. H., & BARRY, J. R. Psychiatric screening of flying personnel: Conditioning and extinction of the galvanic skin response as a function of adjustment to combat crew training. *USAF Sch. Aviat. Med. Rep.*, 1955, No. 55-29.

CATTELL, R. B. Personality structure and measurement: I. The operational determination of trait unities. *Brit. J. Psychol.*, 1946, **36**, 88–103.

CATTELL, R. B. The potentialities of *P* technique deduced from its research applications. Paper presented at Symposium on multivariate analysis of repeated measurements on the same individual, American Psychological Association, Washington, D. C., September 1958.

CATTELL, R. B., & SCHEIER, I. H. *The meaning and measurement of neuroticism and anxiety*. New York: Ronald, 1961.

DEMARIN, F. L., JR. The factor analysis of time series as applied to a problem in cancer chemotherapy. Paper presented at Symposium on multivariate analysis of repeated measurements on the same individual, American Psychological Association, Washington, D. C., September 1958.

HOLTZMAN, W. H., & BITTERMAN, M. E. Psychiatric screening of flying personnel: VI. Anxiety and reactions to stress. *USAF Sch. Aviat. Med. Rep.*, 1952, Proj. No. 21-37-002, Rep. No. 3.

HOLTZMAN, W. H., & BITTERMAN, M. E. A factorial study of adjustment to stress. *J. abnorm. soc. Psychol.*, 1956, **52**, 179–185.

MEFFERD, R. B., JR., MORAN, L. J., & KIMBLE, J. P. Use of a factor analytic technique in the analysis of long term repetitive measurements made upon a single schizophrenic patient. Paper presented at Symposium on multivariate analysis of repeated measurements on the same individual, American Psychological Association, Washington, D. C., September 1958.

MOLNAR, C. E., WEISS, T. F., & GEISLER, C. D. Two techniques for the processing of EEG data. In, *Processing neuroelectric data*. Cambridge: Technology Press of the Massachusetts Institute of Technology, 1959. Ch. 3.

MORAN, L. J. *Repetitive psychological measures*. Austin: University of Texas, Hogg Foundation for Mental Health, 1959.

QUENOUILLE, M. H. *Associated measurements*. New York: Academic Press, 1952.

QUENOUILLE, M. H. *The analysis of multiple time series*. New York: Hafner, 1957.

RESEARCH LABORATORY OF ELECTRONICS, COMMUNICATIONS BIOPHYSICS GROUP, & SIEBERT, W. M. *Processing neuroelectric data*. Cambridge: Technology Press of the Massachusetts Institute of Technology, 1959.

TUCKER, L. R. An inter-battery method of factor analysis. *Psychometrika*, 1958, **23**, 111–136.

(Received February 3, 1961)